



European Association for
Computer-Assisted Language Learning

THE EUROCALL REVIEW

Volume 27, Number 2, September 2019

Editor: Ana Gimeno
Associate Editor: David Perry
ISSN: 1695-2618
<http://polipapers.upv.es/index.php/eurocall>

27

Table of Contents

REVIEW PAPER

A review of MALL: from categories to implementation. The case of Apple's iPad. Valentina Morgana	3
---	---

RESEARCH PAPERS

The effect of online interactions via skype on EFL learners' achievements Yumiko Furumura and Hsin-Chou Huang	13
Assessing Spanish Proficiency of Online Language Learners after Year 1 Rosalie Aldrich and Dianne Money Penny	25

RECOMMENDED RESOURCE

A freely-available system for browser-based Q&A practice in English, with speech recognition Myles O'Brien	35
--	----

TEACHER EDUCATION

Impact and reactions to a blended MA course on Language Education and Technology George S. Ypsilantis	46
--	----

PROJECT

The Linguacuisine Project: A Cooking-based Language Learning Application Paul Seedhouse, Phil Heslop, Ahmed Kharrufa, Simin Ren & Trang Nguyen	65
---	----

Editorial Board	85
-----------------------	----

A review of MALL: from categories to implementation. The case of Apple's iPad

Valentina Morgana
Università Cattolica del Sacro Cuore, Italy

valentina.morgana@unicatt.it

Abstract

The use of mobile devices inside and outside formal settings is often associated with innovative practices in the design of language learning activities. This often implies the reconceptualization of language learning tasks and of the role of the teacher in the mobile classroom. In order to investigate current research and practices in secondary and higher education a review of recent studies in the field of MALL has been undertaken with the main aim of identifying main trends, implementation practices and research gaps.

This paper presents a synthesis of the literature by analysing the four different MALL categories, as presented in Pegrum (2014) and selecting a series of case studies and trends that may be implemented in various educational settings with a specific focus on the use of the iPad in second language settings. The review sought to provide a picture of the various options of MALL task-design and recent implementation practices in secondary and higher education using a specific tablet device. General findings show that many studies are more descriptive than innovative and advocate the implementation of larger and long-term research studies on how mobile devices, and the iPad in particular, are impacting language teaching and learning. Keywords: Mobile-assisted language learning, review, iPad, task design.

1. Introduction

Mobile technologies have many advantages and great potential in affecting and supporting second language learning. These technologies offer an increasingly broad range of devices, such as smartphones, tablets and new generation laptops, implemented by different companies and supporting different operating systems. The rationale behind this review paper is to present theories and studies published in English in the last ten years that examined the use and effectiveness of mobile technologies (specifically tablets and smartphones) and their potential in enhancing language teaching and learning, narrowing the focus on the use of the iPad. The choice of selecting studies specifically oriented to the use of the iOS devices comes from the awareness that different mobile devices (e.g. smartphones, tablets, laptops) have different affordances. The paper looks at current practices and methodologies by presenting single case studies implemented in secondary and higher education. The main aim was to examine the case studies to find valuable intakes to guide future research and practices and to find similar characteristics among iPad focus studies. The paper is broadly divided into four sections: an overview of Mobile Assisted Language Learning categories; recent review studies on the use of the MALL in formal and informal settings; the iPad in language learning contexts and as an inclusive device. The choice of selecting studies specifically designed for iOS devices, such as the iPad, arises from the attempt to find specific features in the various devices and operating systems on the market.

2. Mobile Assisted Language Learning (MALL): the four categories

The term mobile technology is already known in the field of CALL (Computer Assisted Language Learning) (Burston, 2013), but today it goes beyond portable computers such as mobile personal laptops. According to the seminal work of Pegrum (2014) on the use of mobile devices for

languages, literacies and cultures, it is possible to identify different kinds of MALL. In 2009, Garrett (2009) had already applied three different categories to CALL: tutorial CALL, authentic materials-engagement CALL and communication CALL. Pegrum (2014) applied these categories to MALL reorganising them and adding a new one. Following 'a scale of rising (inter)-activity' (Pegrum, 2014 p.94) the categories are: MALL for content, MALL for tutorial, MALL for creation and MALL for communication. The direction of Pegrum's categories clearly moves from a behaviourist approach (content consumption) to a sociocultural approach (content creation and communication). It should be considered that the more mobile devices are integrated in the language classroom the more 'socioculturally informed activities' are possible. In a classroom with a lower level of MALL, content consumption activities (e.g. reading or listening texts) would be the easiest approach to technology enhanced learning and it would start to provide students with autonomous learning activities (e.g. students can listen or read at their own pace). In the tutorial MALL classroom, learners can use different apps to perform a controlled or semi-controlled language task (e.g. podcasts with audio drilling or flashcards). Both MALL for content and MALL for tutorials offer the learners the chance to practise language competences in a low-stress mode since they do not require complex technological skills and the activities are often familiar and similar to those in language books. On the other hand, MALL for creation and MALL for communication require high interactivity of learners with their peers and with the mobile device itself. Moreover, teachers play a key role in the latter two because they provide feedback and support learners in the construction of meaning and knowledge. In creation MALL, students record their texts, take pictures and modify them according to the task, interact with the teachers through specific writing apps, etc. They are actively involved in more sociocultural activities such as collaboration, negotiation of meaning and sharing. Ideally the products created by students in MALL for creation can be shared with peers, teachers, friends and even a wider audience. Many studies in MALL, in fact, show the recent trend to investigate the use of social media in language learning (Mompean & Fouz-González, 2016). In MALL for communication the shift is on the interactive process of activities like reading and writing. Learners can, for example, read collaboratively the same e-books, or write synchronously and asynchronously using sharing services such as Twitter or Instagram.

The studies presented here reflect the wide varieties of MALL uses mainly in the language classroom, although mobile devices allow learners to practise also outside formal settings. Content and tutorial activities, for example, can be moved outside the language classroom allowing the teacher to work on more interactive tasks where support and feedback are required in the classroom.

3. MALL studies in formal and informal learning: attitudes and skills

The use of mobile devices has influenced and is still influencing educational practices and, most importantly, it is creating innovative settings for learning (Pachler et al., 2010). A few studies using various mobile learning devices (smartphones, tablets and laptops) in the fields of science (Lan & Huang, 2012), language courses (Hsu, Hwang, & Chang, 2013), and ICT (Sabah, 2016) suggested that students participate actively in their learning, develop strong collaborative skills, and are able to direct their own learning process.

In their recent review of the developments and implications of MALL studies, Liu et al. (2014) found that 'whenever a new mobile technology is introduced, its effect on language teaching and learning is a popular topic for researchers' (2014, p. 165). Also Hung and Zhang (2012) had already observed the same growing interest in the field of mobile learning in language education in their review of mobile learning studies published between 2003 and 2008. Both studies include a wide range of mobile devices such as smartphones, tablets, Personal Digital Assistants (PDAs) and laptops. These studies indicate the large potential of informal language learning to support formal language activities, showing a need to investigate the interconnection between formal and informal MALL.

In this regard, recent studies have shown that when mobile technologies (smartphones) are integrated effectively in the language classroom, they provide a valuable contribution to the learning approaches, also creating a collaborative learning environment (Sabah, 2016). As such, the literature on mobile technologies for language learning reports a number of case studies that

investigate various aspects and various devices of mobile language learning in formal education (Abdous et al., 2009; Hsu, 2013; Kukulska-Hulme, 2012; Viberg & Grönlund, 2012). In their review of mobile learning, Viberg and Grönlund (2012) observe that the dominating research focus is on the attitudes of learners towards technologies, their intention to use them, and the various uses of mobile technology for authentic communication integrated in their second and foreign language learning. Moreover, most of the recent studies in the field have sustained the idea that mobile technologies can contribute significantly to learners' second and foreign language acquisition in terms of enhancing grammar and vocabulary with smartphones (Çakmak & Erçetin, 2018; Liu, 2016), and developing writing, listening and speaking skills using tablets and smartphones (Chang & Hsu, 2011; Moreno & Vermeulen, 2015; Lin, 2014). Furthermore, a systematic review of MALL studies published from 2008 to 2013 (Liu, Lu, & Lai, 2014) presents evidence of the distribution of learners and language skills (see Figure 1). The review article investigates a total of 24 m-learning articles using the task model for m-learning (Taylor, 2006) in order to identify trends and practices in MALL. Most of the studies examined were conducted with elementary or higher education learners, the mobile devices tested where mainly tablets and smartphones, while reading skills and vocabulary learning proved to be the most investigated areas in language learning.

Year	Levels of learners				Year	Types of language skills					
	Elementary school students	High school students	Higher education	Non-specified		Listening	Speaking	Reading	Writing	Grammar	Vocabulary
2008	0	1	7	0	2008	1	0	2	0	1	7
2009	0	0	2	0	2009	1	1	1	0	0	0
2010	3	0	2	1	2010	2	1	2	1	0	6
2011	2	0	2	0	2011	2	1	1	0	0	1
2012	2	0	2	0	2012	2	1	2	2	0	1
2013	1 ^a	0	0	0	2013	0	0	1	0	0	1
Total	8	1	15	1	Total	8	4	9	3	1	16

Figure 1. Distribution of mobile learners and language skills (from Liu et al, 2014:176).

Based on a study with 45 students from eight different regions and countries in an EFL curriculum with an activity-oriented design, Hsu (2013) found that learners with different cultural backgrounds had varying attitudes towards MALL in terms of technological affordances, applicability and constructivist aspects (e.g. collaborative tasks). However, it was not possible to establish the exact reasons for their different experiences and expectations. According to the survey data, many students did not believe they could practice all language skills in a mobile learning setting, but they may not have had the opportunities to see how it could be effectively done. The review does not take into consideration the difference between the type of mobile devices used in the studies.

Likewise, one of the latest studies observed the attitudes of 345 higher education students in Sweden and China (Viberg & Gronlund, 2013). The researchers found that learners had particularly positive attitudes toward the chance to personalize their learning, the opportunity to have a valid learning experience, and the occasion to exchange information and collaborate with other students, and teachers. The study conducted by Hsu (2013) presents different attitudes of learners towards technology (smartphones) due to their cultural background; and further, Viber & Gronlund (2013) confirm that positive attitudes towards mobile devices can have an important impact in the language classroom. This implies that the context where the studies are carried out necessarily influences the design and implementation of m-learning in language settings. However, one of the studies mentioned above (Hsu, 2013) is quite small, and does not provide clear evidence of learner perceptions and mobile language learning, although it does give an idea of learners' perceptions of smartphones in the language classroom, especially in terms of real communication, personalization of learning, and multimodality (Kukulska-Hulme, 2013).

As stated above, many research studies have focused on the descriptive analysis of language teaching applications to develop specific skills, and in particular on listening, speaking and reading skills. For instance, Chang and Hsu (2011) analysed the use of mobile devices (PDAs) in an intensive reading course with intermediate EFL learners, including the attitudes and satisfaction levels of the users. One of the main ideas of the study was to integrate collaborative learning into

reading activities on a mobile assisted language system by analysing the usage of the system by individual students and by groups of learners. Perceptions and satisfaction around the use of mobile devices were measured using the Technology Acceptance Model (TAM) questionnaire that addresses usefulness and the perceived ease-of-use (Park, Nam, & Cha, 2012). Their study found that the collaborative method gave a meaningful contribution to supporting EFL learners in reading comprehension. Interestingly, students grouped in twos, threes and fours performed better than individual students, but also better than groups of five learners. This implies that collaborative activities performed through mobile technologies, specifically PDAs, can foster language learning in small groups. Moreover, learners liked the mobile device, and they thought it was useful and easy to use to read and annotate texts. However, Chang and Hsu (2011) observed that in addition to enhancing reading comprehension, forthcoming studies should also consider investigating listening, speaking and writing proficiency functions.

The studies presented here offered an overview of the main directions of MALL research considering various types of mobile devices. The following sections will focus on MALL studies designed and implemented using only iOS devices.

4. The case of the iPad in language learning

Language learning is one of the disciplines that could derive many benefits from the use of mobile technologies (Kukulska-Hulme, 2006). Since its first launch onto the market in 2010, the iPad has been implemented in various educational settings. There are several features that often influenced the choice of the Apple device: the screen size (three options available) that resembles the textbook page; the long-lasting battery, the lightweight, the access to a dedicated education App store, a single control button and an on screen keyboard (Albadry, 2017). In the classroom, tablets, such as the iPad, for example, allow learners to record themselves and to listen to audio at any point of the language lesson. Students can be invited to perform authentic interactions, collaborating and creating on their tablet devices. They are also easily exposed to a wide range of authentic materials, which strongly support the integration of language learning with everyday communication needs (Morgana, 2014). The iPad was the first tablet device to provide educators with various working configurations (e.g. with the help of various apps they can provide immediate and personalized feedback to students during lessons or flip the class by projecting any student's work on the screen) and enable learners to perform a wide variety of tasks (Gabarre, Gabarre, Din, Shah, & Karim, 2014).

4.1. Learners' and teachers' perceptions

A group of studies on the use of the iPad in the language classroom focused on learners' and teachers' perceptions (e.g., Gabarre et al., 2014; Wang, Teng, & Chen, 2015). For example, Gabarre et al. (2014), explored how iPads can be used in the language classroom to promote active learning opportunities as in Lys (2013) and Chen (2013). They implemented a qualitative research design in the form of a case study in order to have more detailed insights and understandings of the processes. The study involved one French learner in a Malaysian university. Their findings show that the learner felt comfortable using the iPad in the classroom; she mentioned many ways to use it for educational purposes (YouTube videos, dictionary, immediate search for accurate information on a topic etc.); she did not like to use it for writing activities. Similarly, Wang et al. (2015) observed the implementation of the iPad to support EFL vocabulary acquisition with 74 students in a Taiwanese university. The study provided quantitative and qualitative analysis comparing two data groups and a pre-test/post-test design. The participants were divided into two groups: the experimental group using the vocabulary app on the iPad, and the pen and paper group using the traditional semantic map provided by the teacher to learn vocabulary. Findings show that the experimental group performed better in the post-test. Gabarre et al. (2014) showed that the iPad promoted new and active language learning opportunities, while Wang et al. (2015) implemented a larger study, demonstrating how an iPad app contributed to significant progress in learners' vocabulary acquisition. This implies that, based on the few studies carried out so far, learners' attitudes towards the iPad in the language classroom are positive, and teachers should be encouraged to implement a technological-mediated task design in their classroom. However, there is a lack of details and explanations of the tasks implemented in both studies. In addition, the results cannot be considered conclusive

as they are mostly based on data taken from university students, or small-scale studies (e.g., one student).

4.2. iPads in secondary schools

Although the distribution of iPads in secondary schools is rapidly increasing, especially in the United States (as reported by Bloomberg Business in October 2013 and by an Apple company self-report on iPads in education in 2017), large-scale studies on the iPad in secondary classrooms are still scarce. However, there are a few studies of this nature. For instance, Chou, Block and Jesness (2012) ran a four-month pilot project of one-to-one learning with iPads in four 9th grade classrooms in a large K-12 school district in the United States. They collected the data using three data sources: teacher focus groups, student focus group and classroom observation. The researchers compared notes and collected the main themes which emerged from the data collection (e.g., active engagement, increased time for projects, enhanced teaching with updated information). Their findings showed the positive impact of iPad integration especially in terms of motivation, time management, and digital literacy (digital literacy here means those 'skills to effectively decode and encode meaning in digital channels' as in Pegrum 2014, p. 158). The study also interestingly shows the need to have well-prepared teachers in the classroom, confirming one of the issues raised in various studies that mobile learning activities are not effective if teachers are not comfortable with the technologies being used.

There are a number of studies focusing on the use of the iPad in the English as foreign language secondary classroom (e.g., Lin, 2014; Morgana & Shrestha, 2018; Simpson, Walsh, & Rowsell, 2013). Lin (2014), for example, investigated the effects of using iPads in an Extensive Reading Program on teenage English learners' online activities, reading ability and users' perceptions. Two classes and an English teacher were selected in a senior high school in Taiwan; the study lasted ten weeks; one class was assigned to the mobile group reading on iPads and the other, the PC group, reading on PCs. The researcher triangulated data through the users' learning records, the reading texts, and the Technology Acceptance Model questionnaire. Results showed how the mobile group outperformed the PC group and provided empirical evidence for mobile integration in extensive reading programmes in secondary EFL education.

5. The iPad and the development of speaking skills

We are all aware that the Internet and mobile apps offer a variety of opportunities for language learning listening and speaking practice. Learners can listen to authentic materials (e.g., radio and TV channels, audiobooks), and they can also practice the language through chatting (e.g., Face Time, Skype) or recording their voices (e.g., podcast). Some of these media have recently been investigated as supportive tools for second language learning using iOS devices (Abdous et al., 2009; Ducate & Lomicka, 2009; Gromik, 2012; Lord, 2008; Lys, 2013; Papadima-Sophocleous & Charalambous, 2015; Pegrum, 2014).

Lys (2013) conducted an interesting study in an advanced German class, investigating the integration of the iPad into the classroom and its influence on learners' oral language development. The author particularly focused on how an instructional setting that provides additional conversational opportunities in and outside the classroom with a mobile device (iPad) could impact the quality of students' oral language proficiency. The study was a one-to-one iPad implementation project, and it was part of a larger study at a private American university; it lasted nine weeks, involving 13 students. They were engaged in a variety of speaking, listening and recording tasks. Each week they worked on a scaffolded task, had a real time video chat using Face Time and they had to provide an open-ended recorded speech. Results showed that real-time conversational activities could contribute to advanced learners' speaking proficiency. Students had more time to speak compared to a standard non-iPad class, and they reported a high level of enthusiasm. Various aspects of the study presented by Lys (2013) are relevant for future research (e.g. use of scaffolded activities), although we should also bear in mind some important limitations: the lack of a pen and paper group, the difficulties of assessing speaking performance and the limited number of students involved.

Moreover, there are a number of studies that investigated the use of podcasting to improve students' pronunciation. Some of these found certain improvements (Lord, 2008), others did not (Ducate & Lomicka, 2009). In a study at the University of Cyprus learners used mobile devices (iPod Touch, iPad's brother device) to improve oral reading fluency (Papadima-Sophocleous & Charalambous, 2015). Students recorded themselves reading a text, after practising by following a native speaker model on YouTube. After a content analysis of the data produced by the learners, the researchers found a general improvement in speed and word decoding accuracy. This was probably due to the considerable amount of time that learners spent rehearsing with the mobile device.

The iPad, and mobile devices, in general can also provide unlimited opportunities for fluency-focused speaking production (Pegrum, 2014). For instance, in a study conducted in a Japanese university, students were asked to record a 30-second video on a teacher-selected topic (Gromik, 2012). The author triangulated the video/audio data produced by the students with survey data. Results demonstrated an increasing number of words used by students task after task, and students felt the activities proposed enhanced their oral fluency.

Although the studies presented above show positive results, and generally follow a well-designed approach with a coherent data analysis process, we can argue that some aspects (such as the limited number of students and teachers involved) could limit the reliability of their findings. Additionally, these studies do not provide innovative ideas that can support teachers in the use of mobile devices in the second language classroom. They provide a description of standard and general use of iPads. This shows the need for more research on a wider-ranging use of mobile devices, such as iPads, in MALL.

Despite a large number of studies focusing on the use of the iPad to develop speaking skills (Lys, 2013; Morgana, 2018), but also to enhance vocabulary (e.g. Wang et al., 2015), and reading (e.g. Lin, 2014), studies focusing on grammar learning, pronunciation and writing skills are less represented in the reviewed literature. No study so far has investigated the four skills at the same time (listening, speaking, reading and writing) in the EFL classroom.

6. The iPad as an inclusive m-learning tool

The use of the iPad has also been implemented in primary and secondary institutions globally as an accessible and inclusive m-learning tool (Aronin & Floyd, 2013; Cumming, Strnadova, & Singh, 2014; Flewitt, Kucirkova, & Messer, 2014; Hayhoe, 2013; King et al., 2013; Parsons, 2014; Selner, 2011). A relevant study is the one conducted by Cumming, Strnadova, and Singh (2014) at a private high school in Sydney. The action research study investigated the process and the outcomes of the introduction of the iPad as an inclusive learning tool on teachers and students. The project focused on four students with developmental disabilities attending classes in inclusive settings and five special education teachers. The team of teachers and researchers had bi-weekly meetings to reflect on the practice, they collected students' written assignments, wrote articles on a shared webpage and recorded video interviews with teachers and students. Teachers evaluated learners' outcomes following an inquiry and knowledge building cycle (Timperley et al., 2007 cited in Cumming et al. 2014). They were asked to select learners' and teachers' needs, and based on the results, design tasks for the classroom; at the end of the cycle teachers reflected on the impact of the tasks on learning. Data consisted of interviews, notes from teachers' meetings and classroom observation tables, and they were analysed using an inductive content analysis approach. The study concluded that both students and teachers found the iPads to be motivating and effective tools for learning. In the English classroom in particular, the iPads were used for reading texts, viewing movies and, in general to reduce the time students took to read texts or novels. Although findings were consistent with other studies (e.g., Campigotto, McEwen, and Demmans Epp, 2013), the study has the limitation of being an unrepresentative sample: the sample was relatively small and specialized.

7. Conclusion

In general, the literature reviewed here shows that there is a further need to explore how the use of mobile devices, such as the iPad, can facilitate the development and acquisition of linguistic awareness and language skills, how instructors could engage learners equipped with mobile devices, and how second language tasks that would improve learners' experience could be designed (Ifenthaler & Schweinbenz, 2013). Based on the studies presented above it is possible to identify some of the main characteristics of the use of the iPad in secondary and higher education settings. Collaborative activities based on students' engagement and information sharing appeared to be easy to conduct and provided strong motivation, thus, confirming that the category of MALL for creation is a growing trend in the mobile classroom. Also, some studies demonstrated how the iOS device helped teachers to provide immediate and personalised feedback on the tasks proposed. For instance, teachers and students reported positively on the use of the built-in iPad messenger app to interact inside and outside the classroom (Morgana & Shrestha, 2018). Mobile devices can have similar characteristics and affordances, but the fact that learners and teachers use the same device with the same applications and the same operating system appeared to facilitate scaffolding, instruction flow and the management of various issues such as app selection or sharing features. One of the key aspects often reported in MALL studies is the need for teachers to have a good methodology background and to feel comfortable using the mobile device. All the studies presented above involving the use of the iPad as the only mobile device for the research project, showed a great involvement by the teachers and no issues related to this aspect were mentioned. Apparently, participants feel more comfortable and engaged when using the same device (e.g. the iPad) compared to classrooms where different mobile devices with different operating systems have been selected. This choice also has an impact on task-design and implementation making it simpler for the teacher/researcher. Further research is needed in order to investigate the ways in which the use of specific mobile devices is impacting language learning with secondary school and higher education students. Moreover, none of the studies reviewed here looked at the changes in learners' behaviours throughout the implementation of mobile technologies. The present review also identified a few key areas that are underrepresented in the literature related to the use of iOS devices, in particular the lack of longitudinal and large-scale studies, and of studies focusing on grammar or on writing development using iPads. Generally, this review also reveals the need for replication studies in the field of MALL.

References

- Abdous, M., Camarena, M. M., & Facer, B. R. (2009). MALL Technology: Use of Academic Podcasting in the Foreign Language Classroom. *ReCALL*, 21(1), 76-95.
- Albadry, H. (2015). The effect of iPad assisted language learning on developing EFL students' autonomous language learning. In *Critical CALL—Proceedings of the 2015 EUROCALL Conference*, Padova, Italy (p. 1). Research-publishing.net. DOI: [10.14705/rpnet.2015.000302](https://doi.org/10.14705/rpnet.2015.000302).
- Aronin, S., & Floyd, K. K. (2013). Using an iPad in Inclusive Preschool Classrooms to Introduce STEM Concepts. *Council for Exceptional Children*, 45, 34–39.
- Burston, J. (2013). Mobile assisted language learning: A selected annotated bibliography of implementation studies 1994-2012. *Language Learning & Technology*, 17, 157–225. Retrieved from <http://llt.msu.edu/issues/october2013/burston.pdf>.
- Çakmak, F., & Erçetin, G. (2018). Effects of gloss type on text recall and incidental vocabulary learning in mobile-assisted L2 listening. *ReCALL*, 30(1), 24-47.
- Campigotto, R., McEwen, R., & Demmans Epp, C. (2013). Especially social: Exploring the use of an iOS application in special needs classrooms. *Computers and Education*, 60, 74–86. DOI: [10.1016/j.compedu.2012.08.002](https://doi.org/10.1016/j.compedu.2012.08.002).

- Chen, C.-M., & Chung, C.-J. (2008). Personalized mobile English vocabulary learning system based on item response theory and learning memory cycle. *Computers & Education*, 51, 624–645. DOI: [10.1016/j.compedu.2007.06.011](https://doi.org/10.1016/j.compedu.2007.06.011).
- Chien, Y.-C., & Tsou, V. (2012). Learn English with iPad. Presented at the *International Conference on Digital Content*. National Tainan University.
- Chih-Kai and Hsu, C.-K. C. (2011). A mobile-assisted synchronously collaborative translation–annotation system for English as a foreign language (EFL) reading comprehension. *Computer Assisted Language Learning*, 24(2), 155–180.
- Cumming, T. M., Strnadova, I., & Singh, S. (2014). iPads as instructional tools to enhance learning opportunities for students with developmental disabilities: An action research project. *Action Research*, 12(2), 151–176. DOI: [10.1177/1476750314525480](https://doi.org/10.1177/1476750314525480).
- Ducate, L., & Lomicka, L. (2009). Podcasting: An effective tool for honing language students' pronunciation? *Language Learning & Technology*, 13(3), 66–86.
- Flewitt, R., Kucirkova, N., & Messer, D. (2014). Touching the virtual, touching the real: iPads and enabling literacy for students experiencing disability. *Australian Journal of Language & Literacy*, 37(2), 107–116.
- Gabarre, C., Gabarre, S., Din, R., Shah, P. M., & Karim, A. A. (2014). iPads in the foreign language classroom: A learner's perspective. *3L: Language, Linguistics, Literature*, 20(1), 115–128.
- Garrett, N. (2009). Computer-assisted language learning trends and issues revisited: Integrating innovation. *Modern Language Journal*. DOI: [10.1111/j.1540-4781.2009.00969.x](https://doi.org/10.1111/j.1540-4781.2009.00969.x).
- Greenfield, E. The Implementation of the iPad. In Reading Instruction, Education Masters (2012). Retrieved from http://fisherpub.sjfc.edu/education_ETD_masters/212.
- Gromik, N. A. (2012). Cell phone video recording feature as a language learning tool: A case study. *Computers & Education*, 58(1), 223–230. Retrieved from <http://www.sciencedirect.com/science/article/pii/S0360131511001424>.
- Hayhoe, S. (2013) Accessible, inclusive M-learning: using the iPad as a case study. In: TESOL Arabia (Sharjah Section), Sharjah University Community College. (Unpublished). Retrieved from <http://eprints.lse.ac.uk/52734/>.
- Hsu, C.-K., Hwang, G.-J., & Chang, C.-K. (2013). A personalized recommendation-based mobile learning approach to improving the reading performance of EFL students. *Computers & Education*, 63, 327–336. DOI: [10.1016/j.compedu.2012.12.004](https://doi.org/10.1016/j.compedu.2012.12.004).
- 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.
- Ifenthaler, D., & Schweinbenz, V. (2013). The acceptance of Tablet-PCs in classroom instruction: The teachers' perspectives. *Computers in Human Behavior*. DOI: [10.1016/j.chb.2012.11.004](https://doi.org/10.1016/j.chb.2012.11.004).
- King, a. M., Thomeczek, M., Voreis, G., & Scott, V. (2013). iPad(R) use in children and young adults with Autism Spectrum Disorder: An observational study. *Child Language Teaching and Therapy*, 30(2), 159–173.
- Kukulska-Hulme, A. (2006). Mobile language learning now and in the future. In P. Svensson (Ed.) *Från vision till praktik: Språkutbildning och Informationsteknik (From vision to practice: language*

learning and IT). Sweden: Swedish Net University (Nätuniversitetet), pp. 295–310. Retrieved from <http://oro.open.ac.uk/id/eprint/9542>.

Kukulska-Hulme, A. (2012). Language learning defined by time and place: A framework for next generation designs. In J. E. Díaz-Vera (Ed.), *Left to My Own Devices: Learner Autonomy and Mobile Assisted Language Learning* (pp. 1–13). Emerald Group Publishing Limited.

Kukulska-Hulme, A. (2013). Re-skilling Language Learners for a Mobile World. The International Research Foundation for English Language Education (TIRF), Monterey, USA, pp. 1–16. https://www.tirfonline.org/wp-content/uploads/2013/11/TIRF_MALL_Papers_Kukulska-Hulme.pdf.

Lan, Y. F., & Huang, S. M. (2012). Using Mobile Learning to Improve the Reflection: A Case Study of Traffic Violation. *Educational Technology & Society*, 15(2), 179–193.

Lin, C.-C. (2014). Learning English reading in a mobile-assisted extensive reading program. *Computers & Education*, 78, 48–59. DOI: [10.1016/j.compedu.2014.05.004](https://doi.org/10.1016/j.compedu.2014.05.004).

Lin, C. (2014). Learning English reading in a mobile-assisted extensive reading program. *Computers & Education*, 78, 48–59.

Liu, G.-Z., Lu, H.-C., & Lai, C.-T. (2014). Towards the construction of a field: The developments and implications of mobile assisted language learning (MALL). *Digital Scholarship in the Humanities*. DOI: [10.1093/llc/fqu070](https://doi.org/10.1093/llc/fqu070).

Liu, P.-L. (2016). Mobile English Vocabulary Learning Based on Concept-Mapping Strategy. *Language Learning & Technology*, 20(3), 128–141.

Lord, G. (2008). Podcasting Communities and Second Language Pronunciation. *Foreign Language Annals*, 41(2), 364–379.

Lys, F. (2013). The development of advanced learner oral proficiency using iPads. *Language Learning & Technology*, 17, 94–116.

Meurant, R. C. (2010). The iPad and EFL digital literacy. In *Communications in Computer and Information Science* (Vol. 123 CCIS, pp. 224–234).

Mompean, J. A., & Fouz-González, J. (2016). Twitter-based EFL pronunciation instructions. *Language Learning & Technology*. DOI: [10.125/44451](https://doi.org/10.125/44451).

Morgana, V. (2014). Investigating Students' Perceptions of the Use of the Ipad into the English Language Classroom. In Libreriauniversitaria.it (Ed.), *Conference proceedings. ICT for Language Learning* (p. 258).

Morgana, V., & Shrestha, P. N. (2018). Investigating students' and teachers' perceptions of using the iPad in an Italian English as a foreign language classroom. *International Journal of Computer-Assisted Language Learning and Teaching*, 8(3). DOI: [10.4018/IJCALLT.2018070102](https://doi.org/10.4018/IJCALLT.2018070102).

Pachler, N., Cook, J., Bachmair, B., Kress, G., Seipold, J., Adami, E., & Rummler, K. (2010). *Mobile learning: Structures, agency, practices. Mobile Learning: Structures, Agency, Practices*.

Papadima-Sophocleous, S., & Charalambous, M. (2015, March 20). Impact of iPod Touch-Supported Repeated Reading on the English Oral Reading Fluency of L2 students with Specific Learning Difficulties. *The EuroCALL Review*. Retrieved from <http://polipapers.upv.es/index.php/eurocall/article/view/3639/3871>.

- Park, S. Y., Nam, M.-W., & Cha, S.-B. (2012). University students' behavioral intention to use mobile learning: Evaluating the technology acceptance model. *British Journal of Educational Technology*, 43(4), 592–605. DOI: [10.1111/j.1467-8535.2011.01229.x](https://doi.org/10.1111/j.1467-8535.2011.01229.x).
- Parsons, D. (2014). The future of mobile learning and implications for education. In M. Ally & A. Tsinakos (Eds.), *Increasing Access through Mobile Learning* (pp. 217–229). Commonwealth of Learning and Athabasca University.
- Pegrum, M. (2014). *Mobile Learning: Languages, Literacies and Cultures*. Basingstoke: Palgrave Macmillan.
- Sabah, N. M. (2016). Exploring students' awareness and perceptions: Influencing factors and individual differences driving m-learning adoption. *Computers in Human Behavior*, 65. DOI: [10.1016/j.chb.2016.09.009](https://doi.org/10.1016/j.chb.2016.09.009).
- Sekiguchi, S. (2011). Investigating Effects of the iPad on Japanese EFL Students' Self-Regulated Study. *International Conference "ICT for Language Learning"*, 4–7.
- Selner, A. iPads in the Classroom for Literacy Instruction, Education Masters (2011). Retrieved from http://fisherpub.sjfc.edu/education_ETD_masters/24.
- Simpson, A., Walsh, M., & Rowsell, J. (2013). The digital reading path: researching modes and multidirectionality with iPads. *Literacy*, 47(3), 123–130.
- Taylor, L. (2006). Aspect of teacher-generated language in the language classroom. In Borg, S. (Ed.), *Language teacher research in Europe* (pp. 125–138). Alexandria, VA: TESOL.
- Viberg, O., & Grönlund, A. (2013). Cross-cultural analysis of users' attitudes toward the use of mobile devices in second and foreign language learning in higher education: A case from Sweden and China. *Computers & Education*, 69, 169–180.
- Viberg, O., & Grönlund, Å. (2012). Mobile assisted language learning: A literature review. In *Proceedings of the 11th International Conference on Mobile and Contextual Learning* (pp. 1–8).
- Wang, B. T., Teng, C. W., & Chen, H. T. (2015). Using iPad to Facilitate English Vocabulary Learning. *International Journal of Information and Education Technology*, 5(2), 100–104.
-

The effect of online exchanges via Skype on EFL learners' achievements

Yumiko Furumura* and Hsin-Chou Huang**

*Nagasaki University, Japan | **National Taiwan Ocean University, Keelung, Taiwan

fuyumiko2@live.jp | joehuang@ntou.edu.tw

Abstract

This study examined whether direct communication with people from other countries using Skype or Line would affect students' English test scores in listening and reading as well as the development of their curiosity concerning foreign cultures by comparing the data of an experimental group with that of a control group. The former group conducted online exchanges with foreign students, while the latter group did not. As many Japanese companies engaged in international business require high scores in the TOEIC test, which is one of the multiple-choice English tests of listening and reading often used to show each person's English proficiency, universities in Japan are making efforts to improve their students' scores in such an English test. Preparation classes for English tests have been offered. However, students have been likely to lose interest in learning English in the circumstances of this learning style. Results of the study indicated that although the aim of exchange activities is to foster students' curiosity concerning intercultural matters, students experiencing online exchange with skype significantly raised their scores in TOEIC tests in listening and reading after a programme of synchronous exchanges with foreign students, compared with ones who did not experience such online exchanges.

Keywords: Language teaching methodology, Skype, online exchanges, listening and reading.

1. Introduction

A government's education policies strongly affect its nation's education goals. The Japanese Ministry of Education, Culture, Sports, Science and Technology (MEXT) initiated its Project for Promotion of Global Human Resource Development in 2012. This is a funding project that

aims to overcome the Japanese younger generation's 'inward tendency' and to foster human resources who can positively meet the challenges and succeed in the global field, as the basis for improving Japan's global competitiveness and enhancing the ties between nations. Efforts to promote the internalization of university education in Japan will be given strong, priority support (The Ministry of Education, Culture, Sports, Science & Technology in Japan, 2012).

The 'inward tendency' of Japanese young people refers to the fact that many tend to show interest in matters only inside Japan, and they hope to domestically study and work. There are several reasons for this tendency, including, for example, that they may be satisfied with their present life in a convenient, safe environment.

MEXT reported in December 2010 that the number of Japanese going overseas to study fell in each of the four years 2004–2008, dropping from 82,945 in 2004 to 66,833 at the end of this period (The Japan Times online 2011). In response, MEXT started to encourage many universities in Japan to send their students to foreign universities using the abovementioned funding. The recent interest in preparing Japanese university students to study abroad or work in intercultural contexts has caused a demand for foreign language educators to create effective

intercultural and foreign language development programmes (Miyafusa & Fritz, 2016). Therefore, students must first acquire linguistic competence and intercultural communicative competence (hereafter, ICC) to achieve this purpose (Furumura, 2010). Chi and Suthers (2015: 108) defined ICC as 'the ability to develop meaningful intercultural relations with host and other nationals'. The definition of intercultural communicative competence described by Byram (1997) includes the following five components; Attitudes, Knowledge, Skills of interpreting and relating, Skills of discovery and interaction, Critical cultural awareness. It would be difficult to develop all components at the same time for a short period, but we believe that 'curiosity about other cultures', which is one of the elements of 'Attitude' in ICC, needs to be developed at the first stage of higher education if the MEXT's aim is to be achieved. If Japanese students start to study abroad in the second or third year of university, educators need to encourage them to change their 'inward' attitude by developing their curiosity for international matters, and at the same time improve their linguistic competence when they are in the first year of university.

Educational achievement of each university is evaluated by MEXT in part according to the number of students studying abroad and according to the development of student English test scores for a certain period, such as TOEIC, the Test of English for International Communication developed by the Educational Testing Service (ETS) in the U.S. Many universities in Japan have used the listening and reading section of the TOEIC test to show their students' development of English language skills. The project aims to foster human resources that can positively meet challenges and succeed in the global field. It may appear as a challenge to show evidence of the effectiveness of this type of education. However, English educators are especially expected to increase students' curiosity with regard to cultures and people outside Japan, and to improve scores in English tests such as TOEIC in English education.

Two English language educators met at the international conference on CALL in 2013, and then started to collaborate to teach English to university students in Taiwan and Japan using online exchanges. In the first year, a forum site was used, but this resulted in some students not receiving replies from partner students and becoming disappointed. In the spring semester in 2015, fortunately, classes in the two countries were overlapped for about 45 minutes. Thus, we decided to use Skype for students to synchronously communicate in pairs, i.e., a Taiwanese student matched with a Japanese student. In cases of transmission failures, Line was also used to avoid failures in communication of pairs. A Japanese educator, one of two English language educators engaged in the project, had another class of Japanese students who were unable to have international partners to talk with via skype during the period of their class, where the same contents and activities instructed with the same textbook were taught as the class using skype to communicate with Taiwanese students.

The difference between these two classes of Japanese students was whether they communicated with Taiwanese students in class via Skype or not. Based on a comparison of curiosity development concerning foreign cultures, and English test scores between students in the two classes, the effect of this online communication will be examined.

2. The effects of computer-mediated communication on Linguistic and Intercultural competence

O'Dowd (2016, p.291) states that online learning involves engaging learners in interaction and collaboration with classes in distant locations through online communication technologies. Wikis, as emerging Web 2.0 tools, have been used in language learning (Li, 2012). "A wiki was developed approximately in 1995 as a part of Web 2.0 – the read/write web (p.17)", so it has been used to develop students' skills of reading and writing. Dizon and Thanyawatpokin (2018) also argued that the Web 2.0 tools were found to have equally positive and significant effects on writing fluency and syntactic complexity, while neither CMC method had any effect on lexical richness. In language education, Computer-Mediated Communication (CMC) and telecollaboration using tools such as Wikis have many possibilities for students' development of four skills of a target language and intercultural communicative competence through communication with people by using the target language.

There are two types of CMC, synchronous and asynchronous. Skype and video-conference systems are the former type, which have been used to develop students' speaking skills and pronunciation of their target language (Correa, 2015; Alastuey, 2010; Lee, 2007). Hsu (2015) used text-based synchronous CMC to explore the effect of planning conditions on L2 writing. Godwin-Jones (2013) argued that Video-based language exchanges using Skype or similar teleconferencing tools provide facial expressions resembling face-to face conversations, but some problems may arise from insufficient language skills or lack of knowledge of the other culture. He concluded in the study that "if culture is treated experientially (particularly through direct contact with representatives of another culture) this can have a powerful motivating effect, as students see the practical benefits of increased linguistic and intercultural competence" (p. 9).

Wilden (2007) used two types of CMC, the voice chat and the online forum, for secondary school teachers as part of their in-service teacher training. Objectives of the project included promoting their professional and personal development of their intercultural competence. Wu (2018) studied an intercultural asynchronous computer mediated communication activity between Chinese participants and their American peers and concluded that the success of asynchronous CMC can be attributed to the participants' ability to make sense and make use of discursive practices to negotiate positions and achieve positioning alignment. Thus, CMC can be useful for intercultural studies.

As for asynchronous CMC, Young and West (2018) reviewed 22 peer-reviewed journal articles studying the use of asynchronous multimedia-based oral communication in language learning, and then presented this kind of communication promoted language gains in terms of fluency, accuracy and pronunciation in speaking a target language.

Andresen (2009) also reviewed literature to find out about the components of a successful asynchronous discussion, assessment of it, and the limitations of asynchronous teaching, which is that "learners felt disconnected from the discussions and were left wondering if the experience was actually real" (p. 254). Kol and Schcolnik (2008) reported on student writing in asynchronous text-stimulated forum discussions. Their study, however, showed deep student involvement with the content and with their peers according to their qualitative analysis of students' transcripts.

O'Dowd (2016, p.299) stressed the effectiveness of telecollaborative learning on language and culture skills.

One of the most interesting developments in recent years in the field of telecollaborative learning has been the growth of cross-disciplinary telecollaborative initiatives which engage students not only in "pure" foreign language practice, but also in collaborative projects based on different subject areas. This gives students the opportunity to develop language and culture skills while working on subject content, and it also provides them with different cultural perspectives on the particular subject area.

In summary, CMC in both synchronous and asynchronous types has had positive and significant effects on L2 development. Some studies researched skills of writing, reading and syntactic complexity in language education (Dizon and Thanyawatpokin, 2018; Li, 2012; Hsu, 2015) as well as the development of speaking skills and pronunciation of target languages (Correa, 2015; Alastuey, 2010; Lee, 2007). In addition, other studies focused on intercultural competence improvement using CMC such as Skype, teleconferencing tools, or online forums (Godwin-Jones, 2013; Wilden, 2007; Wu, 2018). However, while studies such as those mentioned above have delved into linguistic and intercultural competence, few have focused on the differences in English test scores between 2 groups, one of which used CMC, and another which did not use CMC in order to show students' development in listening and reading skills in the English language.

If some evidence can show that online interaction with others in language learning can be effective to improve students' scores in listening and reading skills in language tests as well as intercultural competence, more educators might be encouraged to use online interactions to connect their students with others outside classrooms.

3. The present study

As MEXT and many Japanese companies engaged in international business require high TOEIC test scores, universities in Japan are making efforts to improve their students' scores. Preparation classes for English tests such as the TOEIC test have been offered. However, students have been likely to lose interest in learning English in the circumstances of this learning style. This tendency may contradict the purpose of developing student curiosity concerning other countries and cultures where English is used as a lingua franca.

This study focuses on Japanese students' activities using Skype or Line. It was assumed that direct communication with people from other countries would influence the development of their curiosity concerning foreign cultures as well as their English test scores. These hypotheses were examined by comparing the data of an experimental group with that of a control group; the former group had online exchanges with foreign students, but the latter group did not.

Our research questions are as follows:

1. Which group of students, the experimental or control, improved their TOEIC test scores of listening and reading skills after online exchanges with Taiwanese students had occurred?
2. Which group of students, the experimental or control, increased their curiosity concerning foreign countries, cultures, and people more after online exchanges with Taiwanese students had occurred?

4. Methodology

4.1. Participants

Participants were students from three classes: classes A and B of Japanese students, and class C of Taiwanese students. Japanese students from class A communicated with Taiwanese students from class C via Skype or Line, while other Japanese students from class B did not communicate with Taiwanese students. In this study, the development of curiosity concerning international matters and scores of an English test, i.e., TOEIC, were compared between students from classes A and B.

There were 45 first-year students in class A, including 18 females and 27 males, consisting of 41 Japanese, 2 Korean, and 2 Chinese students. In this study, 41 Japanese students from class A were in the experimental group. Class B had 46 students, including 18 females and 28 males; one student was from the second year, but the remaining 45 students were from the first year. All students from both classes belonged to the faculty of Economics at the Japanese University.

The total number of Taiwanese students was 45, including 6 females and 39 males, consisting of 42 first-year students, a second-year student, and 2 fourth-year students. They belonged to the departments of harbor and river engineering, mechanical engineering, electrical engineering, and computer science. All students were engineering majors.

Both classes A and B were in the English Communication I course, which is a compulsory English class for first-year students. This course was held in two time slots, each of which was conducted by half of first-year students in the faculty of Economics. Each time slot had three classes, divided according to placement test scores on mini TOEIC; therefore, the total number of classes in English Communication I was six. Of the three classes, classes A and B were at the lowest level for each time slot. Their TOEIC scores are reviewed in the discussion section. Average score from the sample TOEIC test that Taiwanese students in class C attempted at the beginning of the course was 603; these students were at the intermediate level.

4.2. Course design

Students in classes A and B pursued the same course of study with the exception of class A's communication with Taiwanese students via Skype or Line. They were taught English using the same textbook, which explains other countries' cultural aspects, such as how to name babies, how to greet others, prized possessions, and food and drinks. Some textbook topics were used as discussion topics with Taiwanese students. The textbook also presented activities such as listening to English conversations between native English speakers and people of various countries using English as a foreign language. Because each speaker's English pronunciation is influenced by his/her native language, Japanese students were able to become familiar with various types of English pronunciations. Some activities enabled students to express their opinions about a topic in a unit by practicing it in pairs and groups. Thus, students from class A learned some useful expressions about each topic before communicating about it with their Taiwanese partners. They also wrote about what they would talk about with their partners a week before their communication via skype.

Students from class A wrote about what they had learned while they talked with their Taiwanese partners and how they felt about it in compositions of more than 40 words in English after class as homework. Students from class B also wrote what they had learned in each unit and how they felt about it in written compositions of more than 40 words in English after studying each unit as homework. In addition, students from both classes, A and B, did the same homework of reading and listening using an e-learning system every week. Lessons in both classes consisted of speaking, listening, reading, and writing exercises. The sole difference in the content of lessons between class A and class B was the activity of online exchanges with Taiwanese partners.

Class C was a group of non-English majors taking an intermediate-level freshman English course at a national university in Taiwan. This course was organized around cross-cultural topics that developed all four linguistic skills: listening, reading, speaking, and writing. Teaching materials were selected from web-based resources and were designed by an instructor to enhance students' communication skills. This intercultural exchange project was included in the curriculum to provide students with opportunities to use English for authentic communication. In addition to regular class meetings, students were assigned online research tasks to enrich their knowledge of given topics. After familiarizing themselves with the topics, students communicated with their Japanese peers via Skype to talk about each topic during the first hour of their class meetings. This was the first time students were able to communicate with their global peers to express themselves; it impressed them and motivated them to communicate in English.

Students from classes A and C had six sessions of talking about six topics over the course of the spring semester, when the academic calendars used by universities in Taiwan and Japan overlapped. The university in Taiwan started and finished one and a half months earlier than the university in Japan. Only class A overlapped with class C in Taiwan; class B did not have an equivalent class in Taiwan to communicate with.

Classes A and C had a session of about 30 min to talk about a topic during each class. The topics were as follows:

1. Introduce yourself with an explanation regarding why your parents gave you the name you have, and then, talk about your favorite songs.
2. In your culture, how do you greet one another in face-to-face communication, e.g., bow, hug, or kiss; is it different according to the situation and relationship between people?
3. Talk about your favorite movies.
4. What types of drinks, sweets, snacks, and local food do you like? When do you drink or eat them, e.g., every morning, at a party, before going to bed, or some other time?
5. What days are special to you, e.g., your birthday, New Year holidays, Christmas, or other days? What do you do on these occasions?
6. Talk about your future job and your dream after graduation from your university.

The first five topics, 1–5, were selected from the textbook that Japanese students were studying, and the last topic, 6, was selected by the two educators to allow students to share values about work in their future. A purpose of communication via Skype or Line was to stimulate student curiosity concerning different cultures, particularly the customs and opinions of similar young

people in a different country. They used Skype to synchronously speak with each other. This study focuses on the development of student curiosity and TOEIC test scores as a consequence of students exchanging their opinions via Skype or LINE.

4.3. Procedure

Test scores from mini TOEIC, which students from classes A and B attempted immediately before the course at the beginning of April 2015, were compared by the independent samples t-test. Next, TOEIC test scores obtained by the same students in the latter part of the course, the beginning of July in the same year, were evaluated by the same t-test to address research question 1.

Students evaluated their curiosity concerning foreign countries, cultures, and people before and after the course using a five-point rating scale, ranging from 1 (not interested at all) to 5 (highly interested). Changes in the curiosity of students in each class were examined by the paired samples t-test to determine whether their curiosity changed before and after the course. Then, the differences of points in their rating after the course between the two classes were compared by the independent samples t-test to determine the class of students, class A or B, which increased curiosity more through the course, addressing research question 2.

5. Results

5.1 Research question 1: Which group of students, the experimental or control, improved their TOEIC test scores after online exchanges with Taiwanese students had occurred?

The total number of students in class A was 45, which included 2 Chinese and 2 Korean students. Because this study focuses on the development of Japanese students, these foreign students' data was excluded. In addition, because 2 Japanese students did not attempt the TOEIC test, the total number of subjects in class A for research question 1 was 39. In addition, the total number of students in class B was 46; however, 1 student did not attempt the TOEIC test. Therefore, the total number of subjects in class B for research question 1 was 45.

Means for the mini TOEIC test scores, which the students in the two classes attempted before the course, are shown in Table 1. In the results of the independent samples t-test, there was no significant difference between the levels of the two groups in English proficiency (classes A and B) ($t(82) = .907$, $p = .367$, $d = 0.20$). In other words, in the levels attained on the mini TOEIC test, there was no difference between students from the two classes.

About three months after attempting the mini TOEIC test, students from both classes attempted the TOEIC test at the latter part of the course. The results of the independent samples t-test shown in Table 2 reveal that the scores of the experimental group (class A) ($M = 420.51$, $SD = 79.30$) were significantly higher than those of the control group (class B) ($M = 384.22$, $SD = 81.44$, $t(82) = 2.062$, $p = .042$, $d = 0.45$).

Table 1. Mini TOEIC test scores in the two classes (full mark: 50).

Experimental group (class A)		Control group (class B)		<i>t</i>	<i>p</i>
<i>N</i>	<i>M</i> (<i>SD</i>)	<i>N</i>	<i>M</i> (<i>SD</i>)		
39	39.90 (5.55)	45	38.87 (4.86)	.907	.367

Table 2. TOEIC test scores in the two classes (full mark: 990).

Experimental group (class A)		Control group (class B)		<i>t</i>	<i>p</i>
<i>N</i>	<i>M</i> (SD)	<i>N</i>	<i>M</i> (SD)		
39	420.51 (79.30)	45	384.22 (81.44)	2.062	.042*

$P^* < .05$

5.2. Research question 2: Which group of students, the experimental or control, increased their curiosity concerning foreign countries, cultures, and people after online exchanges with Taiwanese students had occurred?

Students from classes A and B answered the following questions after their completing the course.

1. How much were you interested in foreign countries, their cultures, and their people before the course? Please evaluate it with a five-point rating scale, ranging from 1 (not interested at all) to 5 (highly interested).
2. Please explain the reason for your answer to Q1 in English.
3. How much are you interested in foreign countries, their cultures, and their people after the course? Please evaluate it with a five-point rating scale, ranging from 1 (not interested at all) to 5 (highly interested).
4. Please explain the reason for your answer to Q3 in English.

A within-subject design was adopted to examine how subjects' curiosity changed through the course using a five-point rating scale, ranging from 1 (not interested at all) to 5 (highly interested). A paired samples *t*-test was used to analyse changes. In class A, 40 Japanese students out of 45 answered this question because four students were excluded and one Japanese student was absent from the last class. In class B, 45 students out of 46 answered because one student was absent.

The results of a paired samples *t*-test shown in table 3 on the measures of student curiosity in class A, the experimental group, reveal a significant difference in the students' curiosity in foreign countries, their cultures and people ($t(39) = -8.708$, $p = .000$, $d = .81$), between before the course ($M = 3.23$, $SD = 1.19$) and after the course ($M = 4.65$, $SD = .53$) (see table 3). Cohen's $d = .81$ is considered a large effect size in this result.

Table 3. Change in curiosity before and after in the experimental group (class A).

Curiosity before the course		Curiosity after the course	<i>t</i>	<i>p</i>
<i>N</i>	<i>M</i> (SD)	<i>M</i> (SD)		
40	3.23 (1.19)	4.65 (.53)	-8.708	.000**

$P^{**} < .01$

In the control group (class B), students' curiosity after the course ($M = 4.62$, $SD = .58$) was significantly higher than that before the course ($M = 3.20$, $SD = 1.32$, $t(44) = -7.309$, $p = .000$, $d = .74$) (see the table 4). Cohen's $d = .74$ is considered a large effect size in this result.

Table 4. Change in curiosity before and after in the control group (class B).

Curiosity before the course		Curiosity after the course	<i>t</i>	<i>p</i>
<i>N</i>	<i>M</i> (SD)	<i>M</i> (SD)		
45	3.20 (1.32)	4.62 (.58)	-7.309	.000**

P** < .01

The abovementioned results of classes A and B revealed that the course in both classes achieved the aim of raising students' curiosity concerning international matters, irrespective of whether the course included communication with Taiwanese students or not.

Next, the change in student curiosity before and after the course for the experimental group (class A) was compared with the change for the control group (class B) using the independent samples t-test. The means of the rating points of students from class A were $M = 1.43$ and $SD = 1.03$, and those from class B were $M = 1.42$ and $SD = 1.31$. The calculation of the t-test showed that there was no significant difference in student curiosity change before and after the course for the two classes, A and B, ($t(83) = .011$, $p = .991$, $d = .01$).

Table 5. Comparison of changed curiosity between the two classes.

Experimental group (class A)		Control group (class B)		<i>t</i>	<i>p</i>
<i>N</i>	<i>M</i> (SD)	<i>N</i>	<i>M</i> (SD)		
40	1.43 (1.03)	45	1.42 (1.31)	.011	.991

With this result, educators may be encouraged to increase students' curiosity even if the collaborative or online exchange of opinions with foreign students does not occur. It may sometimes be difficult for some educators to find a partner to communicate with outside the country. Our study found that using only a textbook could strongly influence students' curiosity, which will motivate them to study more.

6. Discussion

6.1. Effect of online exchanges on scores of multiple-choice English tests of listening and reading

Although students in the experimental group obtained significantly higher scores than those in the control group in this study, the real purpose of this course, including online exchanges with foreign students, is not to improve student test scores in tests such as TOEIC but to develop students' ICC, which includes many skills. As people are facing a rapidly globalizing world, the young generation is often instructed by adults such as teachers, parents, and others around them that they need to acquire many skills; using English, discussion, critical thinking, collaborative working, intercultural competences, and the like.

Japanese young people tend to think that they will not communicate or work with foreign people, especially in rural areas of Japan. Recently, however, even small companies producing Japanese sake or Japanese food have begun to export their products to survive in an internationally competitive world. Business situations require negotiation skills and the development of good relations among people involved. In a monolingual country such as Japan, there are few opportunities to communicate with people with different ethnic backgrounds except at sightseeing spots. Therefore, it is difficult for people to naturally develop skills related to ICC in everyday life.

As Byram (2008: 157) states, 'people need certain competences in order to be able to act sensibly in and across political entities, at whatever level'. Japanese young people especially need to acquire (at least some of) the competences included in ICC, the evaluation of which is not commonly established in language education.

Even if educators successfully pursue language education, their effectiveness to persuade the officials of MEXT might fail because of less reliable evidence to convince many people of their successful achievement. However, e.g., increases in scores of English tests or in the number of students studying abroad would be very persuasive for them in evaluating the effectiveness of education. Thus, the significant difference in test scores between the experimental and control groups was emphasized in this study, although the real aim of online exchanges was to develop students' ICC.

Many language teachers may have been under pressure to show evidence to prove their practice in teaching, as Kramsch (2015: 458) argued:

The proliferation of other competences from semiotic competence (van Lier, 2004) to symbolic competence (Kramsch and Whiteside 2008) to intercultural competence (Byram, 1997) to performative competence (Canagarajah, 2014), offered by researchers as a way of preparing language learners for a decentered, global economy, is bewildering for the practitioner who is at the same time under increased pressure to measure and evaluate success through multiple choice tests so as to justify his/her own existence.

The sole difference between the content of lessons for the experimental group and those for the control group was the activity of online exchanges. Because the English level of their Taiwanese partners in the experimental group was higher than that of Japanese students, most wrote in their journals that they were surprised to find that Taiwanese students spoke English very well. Therefore, it could be assumed that for Japanese students, this experience may have motivated them to study harder than students in the control group, who did not experience direct communication with foreign students.

The results from this study would suggest that one of the ways to make students aware of their own real situations is by comparing themselves with others, and then take actions for their development. Many students in the experimental group wrote in their journals that they were not confident in speaking English, and thus, they wanted to develop their skills. Six students, still a small number, out of 41 wrote that they improved their skill in speaking English through the course. Here are the thoughts of one of the students about his experience in the course:

I was interested in communication with Taiwanese student in the class of English Communication. I did my best to communicate my idea in English. I was not good at communicating by English, so I have been not able to communicate with Taiwanese student at the first time. But I carry on with communication at the second time, third time.... I was able to understand hobby, favourite songs and dream of my partner. I experience pleasure of communication. So, I am interested in other culture (Student No.32).

He wrote that he made significant effort to understand his partner and to make his partner understand his ideas by communicating in English. This shows that he began to develop ICC as well as English skills.

6.2. Effect of online exchanges on students' curiosity concerning other countries, their cultures, and their people

Students in the experimental and control groups significantly increased their curiosity concerning other countries, cultures, and people, and there was no significant difference between the degrees of changing curiosity between the two groups before and after the course. In classes A and B, the same textbook was used, which featured young people of various ethnicities speaking of their cultures. Students listened to them speaking English as a lingua franca with different pronunciations, influenced by their native languages. The pronunciations of their English speaking

may have made students imagine countries those speakers were from. As explained in the results of this study, many students increased their curiosity concerning these intercultural matters also by talking about their own culture, which is Japanese culture, in English after practicing the expressions in each unit. In class B, students spoke with their Japanese classmates, while in class A, they spoke or wrote about their ideas related to topics in the textbook with Taiwanese students as well as with their Japanese classmates.

This difference in whom they communicated with did not affect the development of curiosity in both classes. Actually, students in class B had more time to study content in the textbook compared with class A as 30 minutes were spent in communication with Taiwanese students for six class sessions during the semester. This result shows that topics and activities present in the textbook may be sufficiently intriguing to raise student curiosity concerning international matters. Therefore, selecting a textbook or creating content for a class is important in motivating students.

7. Limitation and future directions

It is rare that the class time in the experimental group in Japan overlapped with their partners' class time in Taiwan. This piece of good fortune gave our students very good opportunities for direct communication with each other. Taiwan and Japan have only a one-hour time difference; thus, this exchange worked. Three years ago before this course was conducted, a professor in Canada invited Japanese students to participate in a discussion on a video conference system. However, we were unable to do it because the times of the two classes did not overlap. When Canadian students were having class early in the morning, Japanese students were spending time at home late in the evening.

Many language educators may know about the potential benefit from synchronously connecting their students with foreign students; however, it is difficult to find time to establish such a connection. It could be suggested that students should talk with foreign partners outside their classroom as homework. In fact, students in our study did this to complete the task of talking on a topic in a case where their partner was absent from the class on the day. In this case, educators cannot see or hear their students talking and cannot know how long they continued to talk with their partners. However, when students communicate with their partners asynchronously in or outside the classroom, educators can see their texts messages on line, therefore this way tends to be used easily for educators. In any case, a system to determine suitable partners for exchange may support educators who are still unfamiliar with this type of exchange programme.

In this study, only about 40 students were examined. The number of subjects should be extended to generalize the results of the effect of online exchanges on English as a foreign language achievement.

8. Conclusion

Many researchers have reported the effects of using social networking sites for language learning. In particular, telecollaborative activities have been disseminated among language educators. "These social networking features can maximize students' opportunities for knowledge construction and collaborative language learning" (Liu et al, 2015: 142). Many studies in this area have focused on the development of pronunciation, speaking, and writing in language learning. Few studies have revealed that online exchanges of students' messages and opinions in intercultural situations can influence scores of multiple-choice language tests. The present study has shown evidence that although the aim of online exchange activities is to foster students' curiosity concerning intercultural matters, students experiencing exchange opportunities significantly raised their scores in multiple-choice English tests of listening and reading after experiencing synchronous exchanges with foreign students, compared with others who did not experience such exchanges. The reason for this phenomenon may be that the real communication activities that students were involved in motivated them to study harder than before and that student awareness of the necessity of learning during activities may have led to the effective development of student linguistic competence. Further research to identify the reason for this phenomenon should be pursued in the future.

References

- Andresen, M. A. (2009). Asynchronous discussion forums: success factors, outcomes, assessments, and limitations. *Educational Technology & Society*, 12(1), 249–257.
- Bueno, M. C. (2010). Synchronous-voice computer-mediated communication: Effects on pronunciation. *CALICO Journal*, 28(1), 1.
- Byram, M. (1997). *Teaching and Assessing Intercultural Communicative Competence*. Multilingual Matters.
- Byram, M. (2008). *From Foreign Language Education to Education for Intercultural Citizenship*. Multilingual Matters.
- Canagarajah, A. S. (2014). Theorizing a competence for translingual practice at the contact zone (pp. 78-102). In S. May (ed.), *The Multilingual Turn: Implications for SLA, TESOL, and Bilingual Education*. Routledge.
- Chi, R. & Sutheers, D. (2015). Assessing intercultural communication competence as a relational construct using social network analysis, *International Journal of Intercultural Relations* 48, 108-119.
- Correa, Y.R. (2015). Skype™ conference calls: A way to promote speaking skills in the teaching and learning of English. *PROFILE Issues in Teachers' Professional Development* 17(1), 143-156. Retrieved from <http://dx.doi.org/10.15446/profile.v17n1.41856>.
- Dizon, G. & Thanyawatpokin, B. (2018). Web 2.0 tools in the EFL classroom: Comparing the effects of Facebook and blogs on L2 writing and interaction. *The EUROCALL Review*, 26(1), 29-42.
- Furumura, Y. (2010). Linguistic competence vs. intercultural competence? In Y. Tsai & S.Houghton (eds.), *Becoming intercultural: Inside and outside the classroom* (pp. 125-143). Cambridge Scholars Publishing.
- Godwin-Jones, R. (2013). Integrating intercultural competence into language learning through technology. *Language Learning & Technology*, 17(2), 1-11.
- Hsu, H. C. (2015). The effect of task planning on L2 performance and L2 development in text-based synchronous computer-mediated communication. *Applied Linguistics*, 32, 1-28.
- Kol, S. & Schcolnik, M. (2008). Asynchronous forums in EAP: Assessment issues. *Language Learning & Technology*, 12(2), 49-70.
- Kramsch, C. (2015). A theory of the practice. *Applied Linguistics* 36(4), 454-465.
- Kramsch, C & Whiteside, A. (2008). Language ecology in multilingual settings. Towards a theory of symbolic competence. *Applied Linguistics*, 29(4), 645-671.
- Lee, L. (2007). One-to-one desktop videoconferencing for developing oral skills: Prospects in perspective. *Languages for Intercultural Communication and Education*, 15, 281.
- Li, M. (2012). Use of wikis in second/foreign language classes: A literature review. *CALL-EJ* 13(1), 17-35.

Liu, M., Abe, K., Cao, M., Liu, S., Ok, D. U., Park, J., Parrish, C. M., and Sardegna, V.G. (2015). An analysis of social network websites for language learning: Implications for teaching and learning English as a Second Language. *CALICO Journal* 32(1), 113-152.

Miyafusa, S. & Fritz, R. (2016). Intercultural communicative competence, complexity theory and assessment: Considerations for Creating Effective Intercultural and Foreign Language Development Programs. *Gaku-en*, 910, 1-12.

O'Dowd, R. (2016). Emerging trends and New Directions in Telecollaborative Learning. *CALICO Journal*, 33(3), 291-310. DOI: [10.1558/cj.v33i3.30747](https://doi.org/10.1558/cj.v33i3.30747)

The Japan Times. (2011). Japan far behind in global language of business. *News online*. Retrieved from <http://www.japantimes.co.jp/text/nn20110106f1.html>.

The Ministry of Education, Culture, Sports, Science & Technology in Japan. (2012). Selection for the FY2012 Project for Promotion of Global Human Resource Development. Retrieved from <http://www.mext.go.jp/english/highered/1326675.htm>.

Wilden, E. (2007). Voice Chats in the Intercultural Classroom: The ABC's On-line Project. In R. O'Dowd (ed.), *Online Intercultural Exchange: An Introduction for Foreign Language Teachers* (pp. 269-275). Multilingual Matters.

Wu, Z. (2018). Positioning (mis)aligned: The (un)making of intercultural asynchronous computer-mediated communication. *Language Learning & Technology*, 22(2), 75-94.

Young, E.H. & West, R. E. (2018). Speaking practice outside the classroom; A literature review of asynchronous multimedia-based oral communication in language learning. *The EUROCALL Review*, 26(1), 59-78.

Assessing Spanish Proficiency of Online Language Learners after Year 1

Rosalie S. Aldrich* and Dianne Burke Money Penny**
Indiana University East, USA

*Rosalie.Aldrich@gmail.com | **dburke01@iue.edu

Abstract

Online (OL) second language (L2) courses are becoming more widely offered in the United States; however, little information exists about the effectiveness of OL L2 courses beyond one semester or course. Therefore, the purpose of this study was to assess Spanish students' oral proficiency after completing one year of OL only L2 courses. At the end of year one, students ($n=65$) completed the Versant exam, which scored overall level of oral proficiency as well as four sub-categories: pronunciation, fluency, sentence formation, and vocabulary production. The results showed that 40% of OL Spanish students met the ACTFL benchmark of Intermediate-Low, while 49% scored Novice-High, one level below the benchmark. A portion (15%) of students not reaching Intermediate-Low scored within a few points of the benchmark. A majority of the students also met the benchmark for pronunciation and fluency, but not for sentence formation or vocabulary production. These results show that it is possible for students enrolled exclusively in online Spanish language classes to meet benchmarks. Thus, OL language students can and should be held to the same standards of oral proficiency as their peers in seated classrooms.

Keywords: Spanish; online language learning, ACTFL benchmarks, L2 acquisition/learning.

1. Introduction

As of 2013, approximately 46% of college students have taken an online course (OL), and that statistic continues to grow (Pappas, 2013). The Online Learning Consortium estimates 28% of all students – over 5.8 million – are currently taking an online course (2015). Second language (L2) courses are included in this growing trend. In fact, a 2016 market study forecasts an 8.6% increase specifically in online language course offerings by 2021 (Technavio, 2016). Although offerings of online language are increasing, Blake (2013) noted that most research in the field of Computer Assisted Language Learning has been conducted on the use of technology tools as part of a face-to-face (F2F) or hybrid curriculum, where students complete portions of the course in some combination of F2F and OL modes. The lack of research on the effectiveness of OL only L2 courses is presumably because fewer possible test subjects exist. Zhang echoes, "Research on online language teaching is still in its infancy compared to the rapid growth of the online language teaching practice" (2014, p. 68). With the need for studies focused specifically on OL L2 instruction, this analysis will center on the oral proficiency of students who enrolled exclusively in OL Spanish language courses at a small regional campus in the United States Midwest.

2. Online classrooms

There is some uncertainty about the OL classroom in general. For example, only 29.1% of faculty consider OL learning as effective as F2F (Allen & Seaman, 2016). Similarly, a Gallup poll conducted by *Inside Higher Ed* (2017) found that only 33% of faculty believe learning outcomes can be equivalent between the two modes, but there were large discrepancies between faculty with OL teaching experience and those without (see Table 1).

Table 1. Skepticism in Online Education		
Faculty who agree that online courses are less effective in the...	OL experience	No OL experience
Ability to deliver the necessary content to meet learning objectives	37%	62%
Ability to answer student questions	47%	72%
Interaction with students during class	78%	92%
Interaction with students outside of class	48%	58%
Grading and communicating about grading	14%	32%
Ability to reach «at-risk» students	70%	87%
Ability to reach «exceptional» students	26%	58%
Ability to rigorously engage students in course material	43%	75%
Ability to maintain academic integrity	45%	71%

As Table 1 shows, professors who have OL experience feel less negatively about the OL mode of instruction. While some faculty may have negative perceptions of OL courses in general, studies show they are effective in terms of student engagement (Angelino & Natvig, 2009) and the development of knowledge and skills (Aronoff et al., 2017). Some research suggests that OL classes actually outperform F2F classes, across various disciplines (Angiello, 2010). Research suggests this is also the case for OL L2 classes. In fact, a majority of studies comparing OL and F2F outcomes reinforce the no significant difference phenomenon between the two course modes as described in the seminal study by Russell (1999).

3. Comparing class formats in language teaching

3.1. Computer assisted language learning in the F2F/hybrid classroom

While both traditional F2F and hybrid L2 classrooms include in-person interaction and instruction, use of technology can provide a significant advantage to reaching language learning outcomes (Plonksy & Ziegler, 2016). Technology has long been shown to aid L2 development with an enrichment of input, feedback, and communication (Sauro, 2011; Zhao, 2003). When referencing specific language skill sets (i.e., listening, speaking, reading, writing), Plonksy and Ziegler's (2016) meta-analysis found that technology in the L2 classroom can improve speaking, reading comprehension, vocabulary, grammar, and fluency. Furthermore, when OL tools are incorporated into F2F L2 classrooms, students increase pronunciation skills (Tanner & Landing, 2009), improve vocabulary, sentence formation (Kim, 2014), and learner uptake (Heift, 2010), defined as responding to corrective feedback. Additionally, adding OL paired synchronous chats and wikis to a F2F class, aids in student L2 writing abilities (Oskoz & Elola, 2014).

Beyond skill-based benefits, learner awareness and autonomy have increased when OL tools were incorporated (Guillen, 2014). Additionally, using tools like OL chats can promote an equalization of participation and increased quantity and quality of L2 (Golonka, Bowles, Frank, Richardson, & Freynik, 2014). The learning advantages to text-based and oral OL communication are numerous. These interactions can occur beyond the time constraints of a timed F2F class session, and can provide further benefits by promoting interaction with native speakers, allow more time for comprehension, as well as help develop interlanguage. Further, transcripts of the sessions can be studied after the interaction to master other skills (O'Dowd, 2007). Clearly, these tools as part of a F2F/hybrid curriculum offer unique learning opportunities.

3.2. Online language classes

With regard to developing a strong OL L2 course, those conducted entirely without F2F contact, Sato, Chen, and Jourdain (2017) stress that discipline standards, task-driven design, and extensive use of multimedia platforms should be utilized to meet learning outcomes. Others express the need for extra instructional attention for less engaged, less adapted students, or those

with lower computer literacy levels (Hong & Samimy, 2010; Mahfouz, 2010). When these considerations are accounted for in course design and implementation, OL courses can provide their own unique benefits in language learning and autonomy. For example, Volle (2005) observed significant gains in OL L2 learners' oral proficiency. In addition, independent learning skills can be specially promoted online via task-based instruction (Lee, 2016). It is true that the absence of F2F contact can create challenges, but Hauck and Stickler (2006) found increased L2 production with the OL format. The use of communication via text or video, common teaching strategies in OL classes, provides a sense of support and belonging, reducing isolation and insecurity that can occur OL (O'Dowd, 2007).

OL L2 courses are often examined in comparison to F2F or hybrid courses. These comparisons cover a wide range of skills and attitudes. For example, working in groups in the L2 environment was found to be more successful OL than in the hybrid model (Asoodar, Marandi, Atai, & Vaezi, 2014). English learners separated into hybrid and OL cohorts for a non-credit bearing course scored similarly on achievement tests and a satisfaction survey, but retention was better in the hybrid course (Harker & Koutsantoni, 2005). In general, when comparing OL and F2F modes of L2 learning, some researchers have found the two to be equivalent (Montiel, 2018), and others even found OL to be more effective in reaching learning outcomes (Grgurovic, Chapelle, & Shelley, 2013). When specifically examining oral proficiency, analyses again show equivalence or a small advantage of the OL format. For example, when proficiency was assessed in F2F and OL German courses, results showed comparable proficiency scores for the two groups (Isenberg, 2010). When OL Spanish was compared to F2F Spanish, studies suggest students' overall oral proficiency are not significantly different between the class modes (Blake, Wilson, Cetto, & Pardo-Ballester, 2008; Moneypenny & Aldrich, 2016). OL Japanese students' performance on a simulated oral proficiency interview, when compared to a F2F cohort, was higher in every area (Sato et al., 2017).

While these studies generally show no significant difference between course formats, some researchers point to small sample sizes, poor study methodologies, and ultimately question the validity of OL L2 coursework (Felix, 2008; Zhao, 2003). Others believe that most online programs do not include adequate spoken contact between course members and instructors to promote oral proficiency (Lin & Warschauer, 2015). Furthermore, because of the wide degree of variance between OL L2 courses, from those that offer no synchronous language exchange to more developed offerings with significant interaction (Blake, 2015), it is indeed unwise to generalize the positive results that do exist. In short, our current knowledge in the field of OL L2 is still insufficient.

Much of the research that does exist on OL L2 courses focuses on outcome measures such as a course grade or learner achievement, not on oral proficiency (Van Deusen-Scholl, 2015). Research demonstrating scores on standardized tests and validated measures of proficiency across the spectrum of OL courses is needed (Tarone, 2015), with particular attention to oral proficiency (Blake 2015; Blake et al., 2008). Additionally, very few studies have reported language acquisition in the online setting beyond one semester or one class (Blake et al., 2008; Moneypenny & Aldrich, 2018). Nationally, 70% of students enrolled in Bachelor of Art's degree programs are required to complete two to four semesters of a second language (Lusin, 2012). Therefore, the data reported on one semester or on one class, although informative, limits the scope of knowledge related to the effectiveness of OL L2 courses to meet set educational outcomes and discipline benchmarks.

As demonstrated, there is a lack of research on OL L2 in general, a dearth of research on student proficiency levels in the OL L2 setting, and the need to examine proficiency standards beyond one class/semester; therefore, the following questions are put forth:

- *RQ1: After completing one year of online college Spanish, do students meet the ACTFL benchmarks for overall oral proficiency?*
- *RQ2: After completing one year of online college Spanish, do students meet the ACTFL benchmarks for pronunciation?*
- *RQ3: After completing one year of online college Spanish, do students meet the ACTFL benchmarks for fluency?*

- *RQ4: After completing one year of online college Spanish, do students meet the ACTFL benchmarks for oral sentence formation?*
- *RQ5: After completing one year of online college Spanish, do students meet the ACTFL benchmarks for vocabulary production?*

4. Methods

4.1. Participants

Data were collected at a small regional campus in the United States' Midwest from undergraduate college students ($n=65$) who completed two semesters of first year Spanish language courses exclusively online. All students were required to complete an oral proficiency assessment, conducted by a third party, at the end of the year (i.e., second semester). Consent for their proficiency scores to be included in the study was voluntary and it was not linked to their course grade. This study was approved by the institution's IRB.

A majority of the participants identified as Caucasian ($n=55$, 84.6%) and female ($n=41$, 63.1%). Of the participants, 46 (70.8%) selected that they were not Hispanic/Latino/a. Participants ranged in age from 18 to 46 ($M=25.77$, $SD=7.09$), with a majority of participants ($n=39$, 60%) aged 18 to 23.

4.2. Procedures

Arriba's sixth edition of MySpanishLab was employed for several short audio/video comprehension exercises in the form of multiple-choice, true/false, and short answer questions. Each semester, students were also required to complete five small group conversation sessions with a teaching assistant via Zoom, a video-conferencing program, which lasted approximately half hour each. At the end of the first semester, students completed a one-on-one oral exam with the professor. During the second semester, this exam took place at midterm. Students also used asynchronous computer mediated communication to practice oral skills with three oral composition assignments each semester. Students were given a prompt and had to record a response and post it to the course discussion board in Canvas, a learning management software. The students were allowed to edit and re-record as many times as they desired before posting the final video for grading. Students were also required to comment on two other students' videos.

At the end of the second semester, students took the Pearson's Versant test to assess their oral proficiency skills. The Versant exam is based on the Theory of Automaticity (Cutler, 2003) and Levelt's (1989) Theory of Language Acquisition. The Versant exam correlates ($r=.86$) with the benchmarks developed through the American Council on the Teaching of Foreign Languages (ACTFL, 2016; Pearson, 2011).

Students completed the 15-minute exam over the telephone by responding to 63 questions. Employing a parser and speech recognition, the Versant exam measures two aspects of languages: manner (fluency and pronunciation) and content (sentence mastery and vocabulary) (Fox & Fraser, 2009). The average of these comprises the overall oral proficiency score. To assess pronunciation students read scripted sentences provided to them and repeated words they heard. Pauses, utterances, and words per minute were used to determine fluency scores as students responded to open-ended questions and retold stories. The ability to produce opposites, for example, to a prompt of "up," a student should respond "down," and answer short comparative questions were used to score the vocabulary section. Sentence formation was assessed when the students rearranged words to form sentences with correct syntax in Spanish. After one year of post-secondary language study, the proficiency benchmark set by ACTFL is Intermediate-Low (Versant scores ranging from 33 to 42).

4.3. Data analysis

The demographic and assessment data were analyzed using SPSS. The institution used in this study has multiple online Spanish language instructors who all use a common course shell to deliver the online class. In order to ensure that different instructors did not significantly influence proficiency scores, researchers controlled for this variable in a linear regression. The results indicate that neither the professor for semester one, nor the professor for semester two significantly predicted overall Versant scores, $F(2, 62)=.14$, $p=.87$, $R^2=.004$ (see Table 2).

Table 2. Results from linear regression semester one and semester two instructors

	<i>B</i>	<i>SE B</i>	β
Constant	31.559	2.690	
Professor for 1st semester Spanish	-.014	.283	-.007
Professor for 2nd semester Spanish	.166	.329	.069

5. Results

5.1. Oral Proficiency

Overall Versant scores were examined to answer the first research question, *After completing one year of online college Spanish, do students meet the ACTFL benchmarks for overall oral proficiency?* Versant test scores ranged from 20 to 60 with an average score of 32.60 ($SD=9.70$). Results show that 40% ($n=26$) of students achieved oral proficiency at or above the ACTFL benchmark range of Intermediate-Low (33-42) (see Table 3). However, a majority of the students achieved oral proficiency levels below the ACTFL benchmark ($n=7$, Novice-Mid; $n=32$, Novice-High); It should be noted that nearly one third ($n=10$) of the students below the benchmark were within two points of the Intermediate-Low threshold.

Table 3. Overall Versant results after two semesters

ACTFL Level	Versant Overall Score	Students Scoring in Range
Novice-Mid	20-22	$n = 7$, 10.8%
Novice-High	23-32	$n = 32$, 49%
Intermediate-Low*	33-42	$n = 15$, 23%
Intermediate-Mid	43-52	$n = 8$, 12.3%
Intermediate-High	53-62	$n = 3$, 4.6%
Advanced-Low	63-72	$n = 0$, 0%
*ACTFL Benchmark for end of year 1		

5.2. Pronunciation

Pronunciation scores were analyzed to address RQ2: *After completing one year of online college Spanish, do students meet the ACTFL benchmarks for pronunciation?* Pronunciation scores ranged widely from 26 to 71 ($M=41.98$, $SD=8.41$). A vast majority of the OL Spanish students met or exceeded the ACTFL benchmark for pronunciation ($n=58$, 89%) at the end of the first year of Spanish.

5.3. Fluency

Fluency was addressed in *RQ3: After completing one year of online college Spanish, do students meet the ACTFL benchmarks for fluency?* Students' fluency scores ranged from 20 to 58 with a mean score of 34.63 ($SD=10.50$). At the end of semester two a majority of students ($n=36$, 55.3%) met or exceeded the ACTFL benchmark for fluency. Additionally, six students were two or fewer points from achieving the Intermediate-Low benchmark.

5.4. Sentence formation

The next Versant sub-category, sentence formation, was explored in *RQ4: After completing one year of online college Spanish, do students meet the ACTFL benchmarks for sentence formation?* Sentence formation scores ranged from 20 to 65 with a mean of 30.80 ($SD=12.37$), which is below the ACTFL benchmark. Only 29% ($n=22$) of students met or exceeded the Intermediate-Low threshold. Five students were within two or fewer points from achieving the benchmark in sentence formation.

5.5. Vocabulary

Last, vocabulary was examined in *RQ5: After completing one year of online college Spanish, do students meet the ACTFL benchmarks for vocabulary production?* Vocabulary scores ranged from 20 to 66 ($M=27.89$, $SD=9.49$). A large majority of students ($n=48$, 73.8%) did not meet the vocabulary benchmark score of 33. However, 32% ($n=21$) scored in the Novice-High range, one level below the benchmark.

6. Discussion

The purpose of this study was to assess specific skill sets related to oral proficiency of students who progressed through one year of an exclusively OL language curriculum. Regardless of course format, ACTFL identifies a guidepost for language abilities compared to the length of L2 study. One measure of OL L2 effectiveness is to compare language proficiency results to the ACTFL benchmarks. In this sense, the benchmark proficiency level serves as the control, and allows a comparative assessment of the efficacy of OL language programs.

The results of this study indicate a mixed level of achievement. Pronunciation and fluency skills were mastered to the benchmark level or beyond for a majority of students. However, sentence formation and vocabulary were only mastered by 29% and 16% of online students after one year of study. This indicates that students perform well on skills related to manner and less well on skills related to content, as defined by Fox and Fraser (2009), which is not uncommon in any language course, irrespective of format. For example, González-Lloret and Nielson (2015) employed the Versant as a pre and post assessment of F2F language teaching, and also found lower scores in these two specific areas on both the pre and posttests. Additionally, Money Penny and Aldrich (2018) reported Versant vocabulary mean scores of 29.38 after one year and 35.16 after two years of mixed OL/F2F language study. Researchers often do not report the Versant subscores for the hybrid and OL first year courses; however, they do indicate a first year overall Versant for Spanish with averages varying but similar to the outcomes indicated in this study (Blake, 2008; González-Lloret & Nielson, 2015). Similar overall Versant scores were reported, when F2F and hybrid courses were examined (Isabelli, 2013).

The findings of this study are in line with others, students are meeting the oral proficiency benchmarks with varied levels of success (Blake et al. 2008; Money Penny & Aldrich, 2016). Many of those who did not score to standard were only a few points away from the Intermediate-Low ranking. The natural question arising from these results is how to move more students across the threshold into intermediate proficiency after year one. As pedagogical research shows, greater familiarity with the test formats beforehand would likely aid some students (Jackson & McGlinn, 2000). For example, the assessments in the course do not require that students orally produce

opposites in response to oral prompts and many of the course vocabulary exercises are in written format. This is also the case for sentence formation. As part of coursework, students often write sentences in exercises where they are given the component pieces to construct. However, these prompts are not oral in the course, nor is an oral response required. Besides dedicating more time to practicing vocabulary and sentence formation in general, revising the course so that students practice these skills orally will likely increase familiarity and reduce exam anxiety when students are also asked to perform these tasks on the Versant at the end of the semester.

Increased exposure to L2 is always a good idea in the language classroom. It could be that requiring more than five small group conversation sessions would increase proficiency. However, requiring more synchronous sessions could overload students looking for the flexibility of an OL course. Perhaps addressing the structure of the sessions themselves would be beneficial. Incorporating short warm-up activities where students practice producing opposites and reconstruct sentences in response to oral prompts may provide the practice those students, especially those near the benchmark, need. This is an interesting area for future investigation.

Research conducted on entirely OL language courses is rare. Most studies examine technology use and its effects in hybrid and F2F modes of instruction. In many ways, this is because few institutions offer online only second language instruction, though the number is growing. The assessment data of post-secondary students who have not been exposed to a F2F foreign language classroom as part of their college experience is incredibly valuable because of its scarcity. The results of the current study show that it is possible for students enrolled exclusively in OL Spanish language classes to meet the standards of oral proficiency level benchmarks established by a national professional organization. Thus, OL language students can and should be held to the same standards of oral proficiency as their peers in the F2F classroom.

References

- Allen, I. E., & Seaman, J. (2016). Online report card: Tracking online education in the United States. *Babson Survey Research Group*. <http://www.babson.edu/Academics/centers/blank-center/global-research/Pages/babson-survey-research-group.aspx>
- American Council on the Teaching of Foreign Language. (2016). *Oral Proficiency in the Workplace*. Alexandria, VA: ACTFL Proficiency Guidelines 2012. <https://www.actfl.org/publications/guidelines-and-manuals/actfl-proficiency-guidelines-2012>
- Angelino, L. M., & Natvig, D. (2009). A conceptual model for engagement of the online learner. *Journal of Educators Online*, 6, 1-19.
- Angiello, R. (2010). Study looks at online learning vs. traditional instruction. *The Education Digest: Essential readings condensed for quick review*, 76, 56-59.
- Aronoff, N., Stellrecht, E., Lyons, A. G., Zafron, M. L., Glogowski, M., Grabowski, J., & Ohtake, P. J. (2017). Teaching evidence-based practice principles to prepare health professions students for an interpersonal learning experience. *Journal of the Medical Library Association*, 105, 376-384. doi.org/10.5195/jmla.2017.179
- Asoodar, M., Marandi, S. S., Atai, M. R., & Vaezi, S. (2014). Learner reflections in virtual vs. blended EAP classes. *Computers in Human Behavior*, 41, 533-543. doi.org/10.1016/j.chb.2014.09.050
- Blake, R. (2013). *Brave new digital classroom: Technology and foreign language learning*. Washington, DC: Georgetown University Press.
- Blake, R. (2015). The messy task of evaluating proficiency in online language courses. *The Modern Language Journal*, 99, 408-412. doi.org/10.1111/modl.12234 5

- Blake, R., Wilson, N. L., Cetto, M., & Pardo-Ballester, C. (2008). Measuring oral proficiency in distance, face-to-face, and blended classrooms. *Language Learning & Technology*, 12, 114-127.
- Cutler, A. (2003). Lexical access. In L. Nadel (Ed.), *Encyclopedia of cognitive science* (Vol. 2), *Epilepsy – Mental imagery, philosophical issues about* (pp. 858-864). London: Nature Publishing Group.
- Felix, U. (2008). The unreasonable effectiveness of CALL: What have we learned in two decades of research? *ReCALL*, 20, 141-161.
- Fox, J., & Fraser, W. (2009). Test review: The Versant Spanish Test. *Language Testing*, 26, 313-322. doi.org/10.1177/02655322090260020602
- Golonka, E. M., Bowles, A. R., Frank, V. M., Richardson, D. L., & Freynik, S. (2014). Technologies for foreign language learning: a review of technology types and their effectiveness. *Computer Assisted Language Learning*, 27, 70-105. doi.org/10.1080/09588221.2012.700315
- González-Lloret, M., & Nielson, K. B. (2015). Evaluating TBLT: The case of a task-based Spanish program. *Language Teaching Research*, 19, 525-549. doi.org/10.1177/1362168814541745
- Grgurović, M., Chapelle, C. A., & Shelley, M. C. (2013). A meta-analysis of effectiveness studies on computer technology-supported language learning. *ReCALL*, 25, 165-198. doi.org/10.1017/S0958344013000013
- Guillén, G. (2014). Expanding the language classroom: Linguistic gains and learning opportunities through e-tandems and social networks. Dissertation. UC Davis, Davis, CA.
- Harker, M., & Koutsantoni, D. (2005). Can it be as effective? Distance versus blended learning in a web-based EAP programme. *ReCALL*, 17, 197-216. doi.org/10.1017/S095834400500042X
- Hauck, M., & Stickler, U. (2006). What does it take to teach online? *CALICO*, 23, 463-475. [doi:10.1558/cj.v23i3.463-475](https://doi.org/10.1558/cj.v23i3.463-475)
- Heift, T. (2010). Prompting in CALL: A longitudinal study of learner uptake. *Modern Language Journal*, 94, 198-216. doi.org/10.1111/j.1540-4781.2010.01017.x
- Hong, K. H., & Samimy, K. K. (2010). The influence of L2 teachers' use of CALL modes on language learners' reactions to blended learning. *CALICO*, 27, 328.
- Inside Higher Ed (2017). *Survey of Faculty Attitudes on Technology*. <https://www.insidehighered.com/booklet/2017-survey-faculty-attitudes-technology>
- Isabelli, C. A. (2013). *Student learning outcomes in hybrid and face-to-face beginning Spanish language courses*. Paper presented at The Future of Education. Florence, Italy. Retrieved from <https://conference.pixel-online.net/FOE/conferenceproceedings.php>
- Isenberg, N.A. (2010). *A comparative study of developmental outcomes in web-based and classroom-based German language education at the post-secondary level: Vocabulary, grammar, language processing, and oral proficiency development* (Doctoral dissertation). (UMI. 3420155).
- Jackson, E. W., & McGlinn, S. (2000). Know the test: One component of test preparation. *Journal of College Reading and Learning*, 31, 84-93. doi.org/10.1080/10790195.2000.10850104
- Kim, S. (2014). Developing autonomous learning for oral proficiency using digital storytelling. *Language Learning & Technology*, 18, 20-35.

- Lee, L. (2016). Autonomous learning through task-based instruction in fully online language courses. *Language Learning & Technology*, 20, 81-97.
- Levelt, W. (1989). *Speaking: From intention to articulation*. Cambridge, MA: MIT Press.
- Lin, C., & Warschauer, M. (2015). Online foreign language education: What are the proficiency outcomes? *The Modern Language Journal*, 99, 394-397. doi.org/10.1111/modl.12234 1
- Lusin, N. (2012). The MLA survey of postsecondary entrance and degree requirements for languages other than English, 2009-10. New York: Modern Language Association. https://www.mla.org/content/download/3316/81618/requirements_survey_200910.pdf
- Mahfouz, S. M. (2010). A study of Jordanian university students' perceptions of using email exchanges with native English keypals for improving their writing competency. *CALICO*, 27, 393-408. doi.org/10.1080/09588220902920151
- Money Penny, D., & Aldrich R. S. (2016). Online and face-to-face language learning: A comparative analysis of oral proficiency in introductory Spanish. *Journal of Educators Online* 13 (2), 105-133.
- Money Penny, D., & Aldrich, R. S. (2018). Developing oral proficiency in Spanish across class modalities. *CALICO: Computer-Assisted Language Instruction Consortium*, 35, 257-273. doi:10.1558/cj.34094.
- Montiel, M. L. (2018). *Comparing online English language learning and face-to-face English language learning at El Bosque University in Colombia*. Richmond, VA: Virginia Commonwealth University.
- O'Dowd, R. (2007). *Online intercultural exchange: An introduction for foreign language teachers*. Clevedon, UK: Multilingual Matters.
- Online Learning Consortium. (2015). Online report card: Tracking online education in the United States. <https://onlinelearningconsortium.org/read/online-report-card-tracking-online-education-united-states-2015/>
- Oskoz, A., & Elola, I. (2014). Promoting foreign language collaborative writing through the use of Web 2.0 tools. In González-Lloret, M. & Ortega L. (eds.), *Technology and tasks: Exploring technology-mediated TBLT*. New York: John Benjamins, 115-148. doi.org/10.1075/tblt.6.05osk
- Pappas, C. (2013). Top 10 e-learning statistics for 2014 that you need to know. <https://elearningindustry.com/top-10-e-learning-statistics-for-2014-you-need-to-know>.
- Pearson (2011). Versant™ Spanish Test. Test description and validation summary. <http://www.versanttest.com/technology/VersantSpanishTestValidation.pdf>
- Plonsky, L., & Ziegler, N. (2016). The CALL-SLA interface: Insights from a second-order synthesis. *Language Learning & Technology*, 20, 17-37.
- Russell, T. (1999). *The no significant difference phenomenon*. Chapel Hill, NC: Office of Instructional Telecommunications, University of North Carolina.
- Sato, E., Chen, J. C. C., & Jourdain, S. (2017). Integrating digital technology in an intensive, fully online college course for Japanese beginning learners: A standards-based, performance-driven approach. *The Modern Language Journal*, 101, 756-775.

Sauro, S. (2011). SCMC for SLA: A research synthesis. *CALICO*, 28, 369–391.
doi.org/10.11139/cj.28.2.369-391

Tanner, M. W., & Landing, M. L. (2009). The effects of computer-assisted pronunciation readings on ESL learners' use of pausing, stress, intonation, and overall comprehensibility. *Language Learning & Technology*, 13, 51-65.

Tarone, E. (2015). Point-counter point measuring proficiency outcomes in online foreign language education. *The Modern Language Journal*, 99, 633-634.

Technavio (2016). *Online language learning market in the US 2017-2021*.
<https://www.technavio.com/report/>

Van Deusen-Scholl, N. (2015). Assessing outcomes in online foreign language education. *The Modern Languages Journal*, 99, 398-400.

Volle, L. (2005). Analyzing oral skills in voice e-mail and online interviews. *Language Learning & Technology*, 9, 146-163.

Zhang, S. (2014). An evidence based guide to designing and developing Chinese-as-a-foreign-language (CFL) courses online. *International Journal of Technology in Teaching and Learning*, 10, 52–71.

Zhao, Y. (2003). Recent developments in technology and language learning: A literature review and meta-analysis. *CALICO*, 21, 7-27.

A freely-available system for browser-based Q&A practice in English, with speech recognition

Myles O'Brien
Mie Prefectural College of Nursing, Japan

myles.obrien@mcn.ac.jp

Abstract

A browser-based system to facilitate practice in asking and answering simple questions in English was developed. The user may ask or answer by speaking or typing, and the computer's output is in the form of speech and / or text. The types of questions handled and the permitted vocabulary are limited, though the vocabulary items may be edited freely. The system was well received in a small pilot study among Japanese students. It is freely-available for download, and requires no technical expertise to deploy, just the facilities and ability to edit text files and upload to the internet.

Keywords: Automatic speech recognition, computer-assisted language learning, Google speech API, pattern practice.

1. Introduction

The system is named AARDVARK (Audibly Ask and Respond to Dynamic Vocabulary Addition and Removal Knowledgebase). It aims to provide a stimulating environment for learners of English to practice asking and answering simple questions. It has much in common with a system developed by the author in 1997 (O'Brien, 1997) but takes advantage of the huge developments in internet technology since those days. The old system was written in HyperCard, a very popular CALL tool at that time, so it could run only on Macintosh computers as a standalone program. It employed the very mechanical-sounding speech simulation available then, and had no speech recognition capability. The current system was written with HTML5 and JavaScript, so it will work in any modern browser. The speech recognition function uses Google's automatic speech recognition (ASR) facility, which works only in the Chrome browser, though not on iOS. So, the system can be fully used on Windows, Mac OS, Linux, and Android, and with all features except speech recognition on iOS. A previous paper (O'Brien, 2017) has given a brief history of the use of speech recognition in CALL, and other accounts are given elsewhere (van Doremalen, Boves, Colpaert, Cucchiari, & Strik, 2016; Ashwell & Elam, 2017; Daniels & Iwago, 2017).

The system includes a sample set of vocabulary items, so it can function out of the box. The items may be edited freely by the user, and the system constructs the questions and answers it generates based purely on its supplied vocabulary and the rules of English grammar and syntax, as written into its algorithms. Its choices are random, without any element of artificial intelligence. Therefore, a judicious choice of vocabulary is recommended to better simulate meaningful interaction. How the system fits into the overall scheme of CALL is considered in the Discussion section.

The system contains two complementary parts, one which answers the user's questions, and the other which asks questions and checks the user's answers. Since the question forms and vocabulary items which may be used are predefined, they are quite limited, though it was attempted to enable a good selection of basic Q&A types for pattern practice.

2. The system interfaces

First, the basic functioning of the system will be outlined through its interfaces. The system consists of two separate, complementary parts:

1. “q-answer” which answers questions put by the user
2. “q-ask” which poses questions to the user and checks the answer given

“q-answer” has 3 interfaces. Figure 1 shows the initial setup screen, which comes up at the start. This allows you to check the available vocabulary, and edit it if you wish. Clicking on “OK” takes you to the main interface.

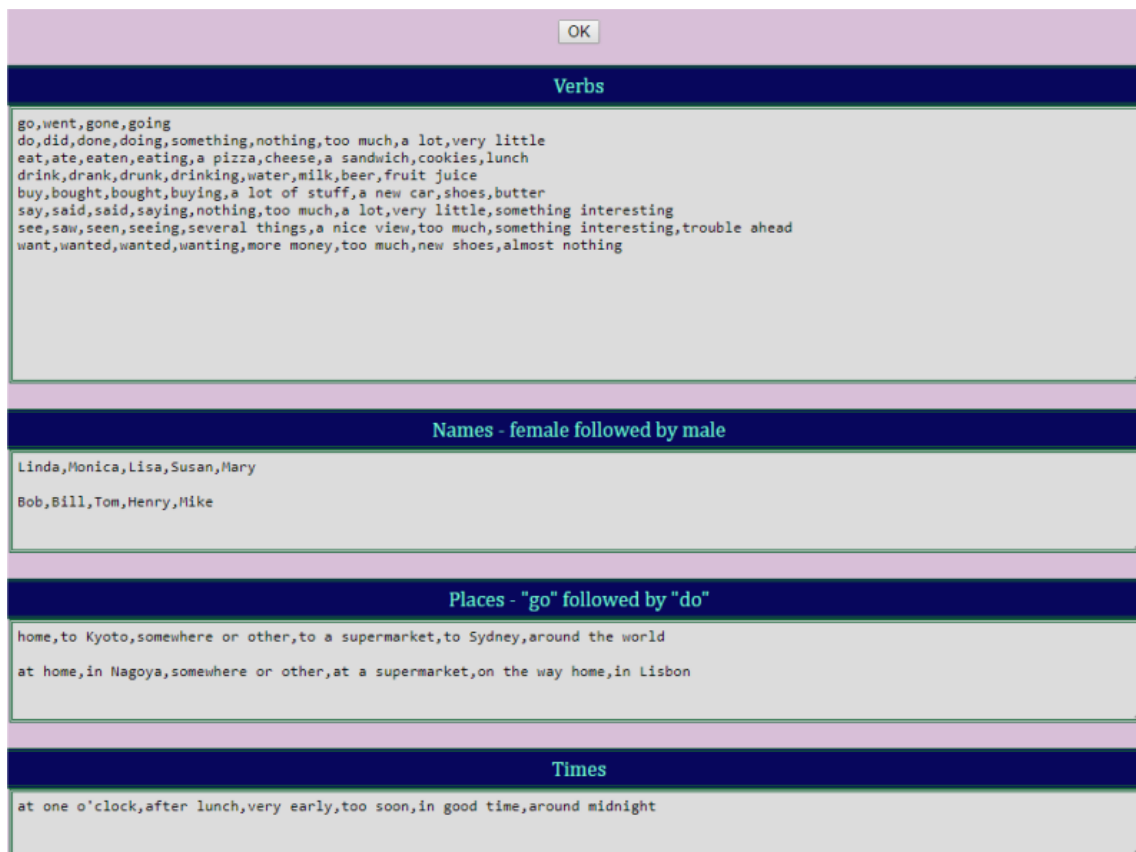


Figure 1. Initial screen for “q-answer”.

Figure 2 shows part of the main screen. You can type a question directly, or click the microphone icon for voice input. Clicking “Enter” calls up an answer. Clicking on the speaker icon allows you to hear the answer again. There is also a tools icon to adjust settings.

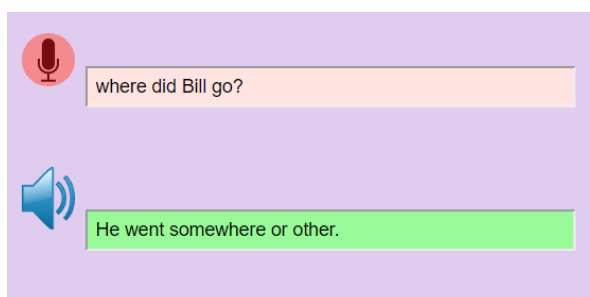


Figure 2. Main screen for “q-answer”.

Figure 3 shows the settings screen which appears when the tools icon is clicked. “Edit vocabulary” brings you back to the initial screen again. “Toggle sound” turns off/on the simulated speech reading of the answer, and “Toggle text” turns off/on the text display of the answer.

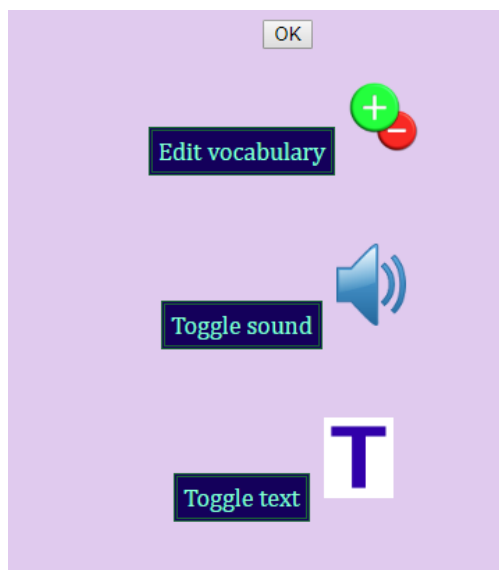


Figure 3. Settings screen for “q-answer”.

“q-ask” also has 3 interfaces. Figure 4 shows the initial setup screen, which comes up at the start. This allows you to check the available vocabulary and the checked question type options, and edit it if you wish. Clicking on both “OK” allows you to accept and proceed to the main interface.

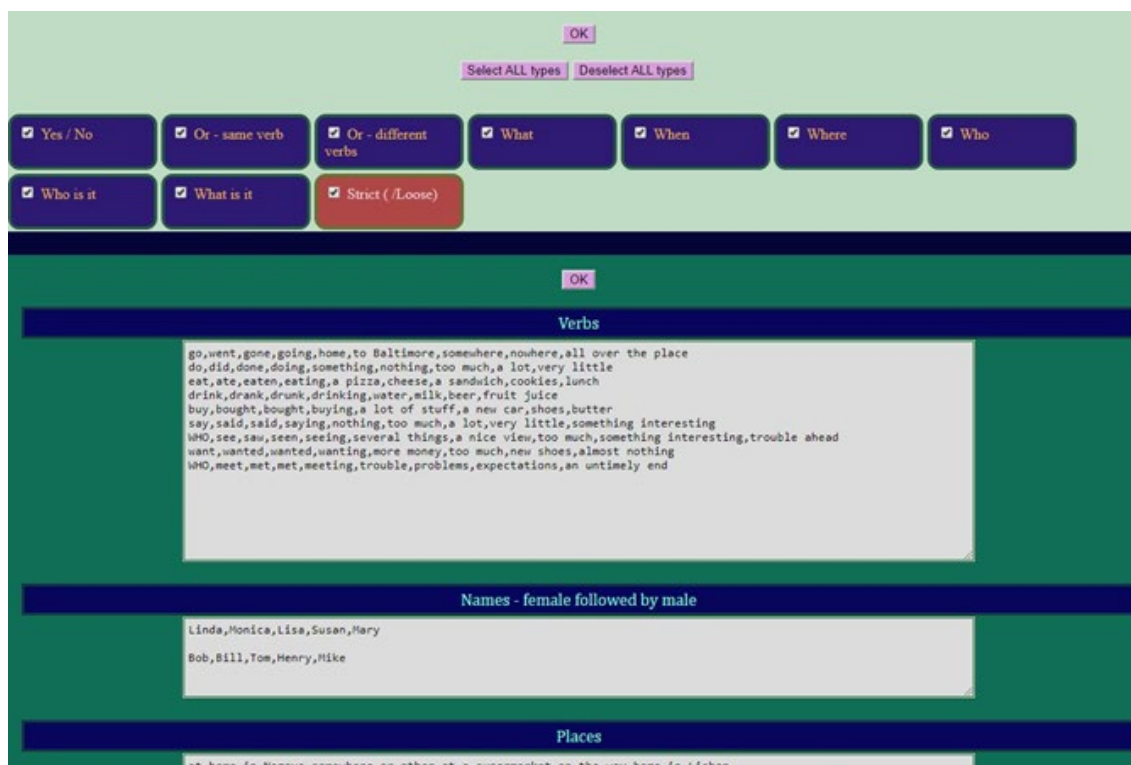


Figure 4. Initial screen for “q-ask”.

Figure 5 shows the main screen. Clicking on “New” generates a question. The microphone icon allows you to speak your answer, or else you can type it. The “Check” button allows you to check your answer. If incorrect, a correct example will be shown. Selecting “Cheat” displays the correct example before inputting your answer. By clicking on the speaker icon, you can hear the question again. Lastly, clicking on the tools icon allows you to adjust settings.

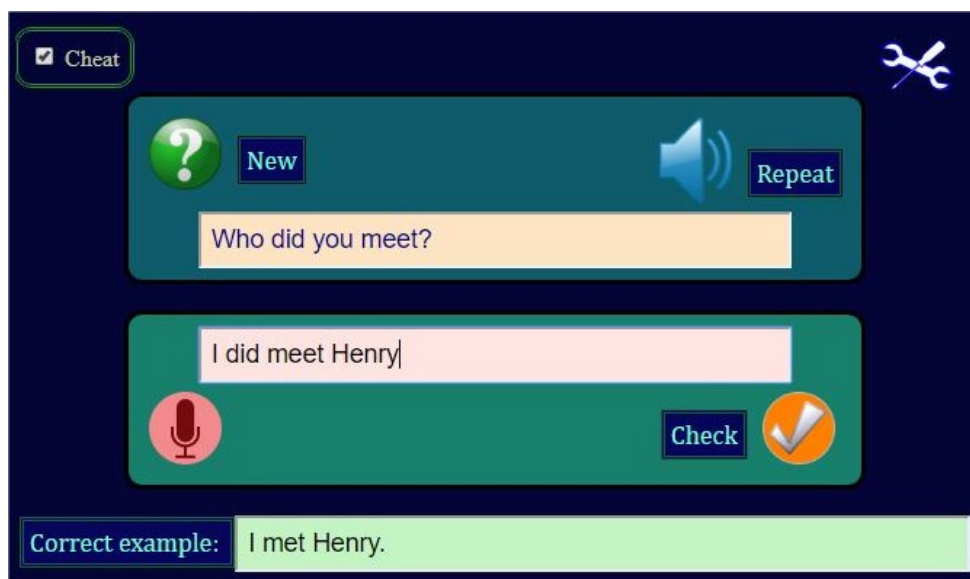


Figure 5. Main screen for “q-ask”.

Figure 6 shows the settings screen from clicking the tools icon. “Edit vocabulary” or “Edit question types” brings you to the setup screen again. “Toggle sound” turns off/on the simulated speech reading of the question, and “Toggle text” turns off/on the text display of the question.

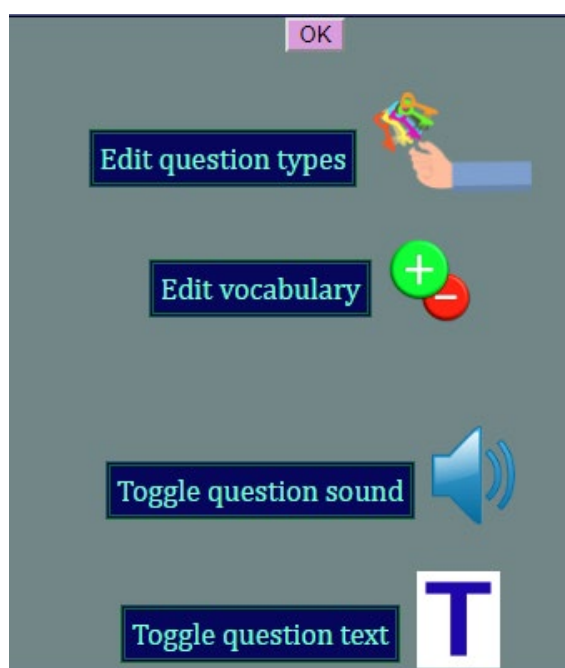


Figure 6. Settings screen for “q-ask”.

3. Scope of the system

3.1. “q-answer”

The user’s question is first analyzed to see which pattern it fits, and the relevant algorithm is triggered to produce a suitable answer. Answers are chosen randomly from the possible alternatives. The types of questions handled, with some sample answers produced when using the vocabulary shown in Figure 1, are as follows:

◆ yes / no questions

- Did Tom see Chris? *No, he didn’t.*
- Will you help me, please? *Yes, I will.*
- Should Linda buy a cat? *No, she shouldn’t.*
- Am I annoying you? *Yes, you are.*

Note that only the first two words, the auxiliary and subject, are relevant in determining the answer, so words outside the specified vocabulary, including nonsense words, may be used in the rest of the question.

◆ either / or questions

- Did you drink tea or coffee? *I drank coffee.*
- Will Linda drink tea or eat chocolate? *She’ll eat chocolate.*
- Should he eat a pizza or something healthy? *He should eat something healthy.*
- Is Mike going today or next week? *He’s going today.*

Again, words outside the specified vocabulary may be used, but the verbs must be specified ones to elicit a proper response.

◆ what questions

- What were you eating? *I was eating lunch.*
- What did Bill buy? *He bought shoes.*
- What did Mary say? *She said very little.*
- What must Linda eat? *She must eat a sandwich.*

In this case, the algorithm will choose a random item from the list of objects shown in the initial screen.

◆ when questions

- When did Lisa eat the cake? *She ate it at one o’clock.*
- When will Mike go to France? *He’ll go there after lunch.*
- When should I eat the grapes? *You should eat them in the morning.*
- When were you drinking coffee? *I was drinking it on Monday.*

In this case, the algorithm will choose a random time from the set-up list, so it is important to include only items which will give a sensible answer regardless of tense. E.g., “next week” might give rise to an answer like “She went there next week.”

◆ where questions

with “go”:

- Where did Mary go yesterday? *She went to Kyoto.*
- Where were you going? *I was going home.*

other:

- Where did Susan eat? *She ate somewhere or other.*
- Where can Mike buy the fish? *He can buy it in Lisbon.*

The “places” vocabulary list is divided into “go” and “do”, so that the vital basic verb “go” can be handled properly.

◆ *who* questions

Name as subject:

- Who did Tom see? *He saw Mary.*

Name as object:

- Who wanted to see Linda? *Mike wanted to see her.*

No name:

- Who was eating the cake? *Tom was eating it.*
- Who is (was) it that...:
- Who was it that ate the cake? *It was Monica.*

With the above types, it was hoped to facilitate practice of a wide variety of examples of question and answer forms in a stimulating way, since the learner asks their own questions, albeit within a limited framework.

3.2. “q-ask”

Although “q-answer” provides a type of interactive practice with question and answer, it is all one-way traffic with the learner passively receiving the answers and leaving the tricky grammatical transformations to the computer. This may be useful up to a point, but more complete practice of communication is desirable, so “q-ask” reverses the roles. It handles almost exactly the same range of types as “q-answer”, though the “name as object” and “no name” types of “who” questions have not yet been implemented, and a “What is it that...” (analogous to “Who is it that...”) type has been added.

The number of options on the initial screen is also much greater with “q-ask”. With “q-answer”, the user decides what questions to ask, so practice is automatically tailored to whatever structures it is desired to use. The default mode of “q-ask” is random, so that what question type will be asked each time is unpredictable. But the initial settings allow any subset of types to be selected, and the questions will be limited to those types. There is also a choice of “Strict” or “Loose” mode. The default Strict mode enforces practice with designated vocabulary items, whereas Loose mode allows greater freedom. With Strict mode, items from the listed vocabulary must be used for objects, names, places and times. Loose mode checks only the basic sentence structure and grammatical transformations. So, taking the question “What did Mike eat?” as an example, “He ate a sandwich.” will be acceptable in either mode if “a sandwich” is one of the objects listed after “eat”. If “a sandwich” is not listed, Loose mode will accept the answer, but Strict mode won’t. Both modes will reject “He eat a sandwich”. Loose mode will also accept the likes of “He ate GBtt%\$fqp skGgm” because the basic structure is correct, despite the nonsensical object. Thus, the learner is afforded flexible communicative practice in that, not alone are the vocabulary lists editable, but items outside the listed ones may be used.

The speech recognition capability is another aspect of the system which contributes to its communicative character. As in the system described in a previous paper (O'Brien, 2017), speech recognition was implemented using Google speech recognition. This is the only ASR engine that offers a web-based API, so it can be easily integrated into a browser-based system like the present one. Furthermore, in comparisons with other systems, it has demonstrated the most accurate speech recognition results (Ashwell & Elam, 2017; Daniels & Iwago, 2017; Kępuska & Bohouta, 2017).

4. File structure of the system

“q-answer” and “q-ask” are separate systems, but almost identical in structure. They consist of a single HTML file, a folder containing images for the interface elements, and four text files containing vocabulary items, one file each for verbs, names, times, and places. Figure 7 depicts the structure diagrammatically.

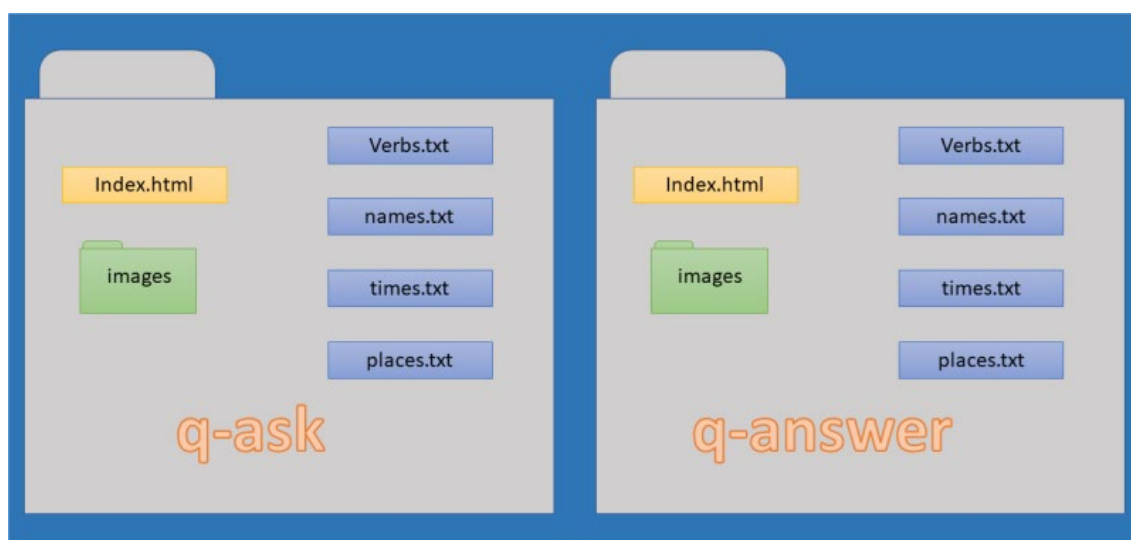


Figure 7. File structure of the entire system.

Both parts are freely available for download. Each one is in the form of a zip file, which expands into a folder as shown in Figure 7. If the folder is uploaded to a website, it should operate immediately without any need for setup or adjustment. In many cases, the uploader may want to edit the vocabulary items, rather than just using the default ones supplied. This is easy to do just by editing the comma-separated lists in the supplied plain text files. The user can also edit the items online while using the system, but the changes will not persist after the browser window is closed. The image files for the interface elements are in the “images” folder, and may be replaced by alternative images if the same filename is used.

An important point to note is that, for speech recognition to work, HTTPS must be enabled. HTTPS support is becoming more common as a free option on hosting services. If it is not available by default, an SSL certificate to enable it can be obtained from a Certificate Authority such as Let's Encrypt, <https://letsencrypt.org>, which is a non-commercial organization offering a free, automated service.

5. Pilot study

A small pilot study to assess user reaction was carried out with ten Japanese nursing students, of mixed English ability. The functioning of both parts of the system was explained and demonstrated to two students at a time, and then they were asked to try out both parts. First one student asked 5 or 6 questions to the “q-answer” system, using voice input, then the other student did the same. The same procedure was repeated once, before moving on to answering questions asked by the “q-ask” system in a similar, alternating way. The students sat beside each other and

communicated freely during the process. The reason for doing the trial in pairs was that it had been noticed before in informal testing with individual students that they found it quite mentally tiring to practice for more than a few questions (especially when answering rather than asking) while being observed by a teacher. Practicing in pairs was found to make the process less stressful and tiring, as it reduced the individual workload and lightened the atmosphere, facilitating mutual help and encouragement.

After the trial session, they were asked to complete a short web questionnaire in Japanese. In English translation, the survey questions asked which activity, asking or answering questions, they thought most likely to be helpful to their English study ("Can't say" was a third alternative). Then, about the system in general, there were three 5-point scale items, "Is likely to help your study of English?", "Is easy to use?", "Is enjoyable to use?", and three free input items, "Good points", "Points needing improvement", "Any other comments". While the researcher observed their use of the system, providing assistance if needed, the questionnaires were completed anonymously in private, to avoid causing pressure.

The responses for the 5-point scale items were quite positive: an average of 3.9 for "Is easy to use", 4.4 for "Is likely to help your study of English?", and 4.8 for "Is enjoyable to use?" In the free input section, positive points mentioned were, "good for pronunciation training" (most frequent), "makes one careful about singular/plural, articles & pronouns", "conversation-like practice", "good for enjoying study with a friend", and "getting correct answer is enjoyable and builds confidence". Among the negative points, by far the most common was the difficulty of getting one's intended spoken words recognized correctly. Suggestions for improvement included having sound as well as text for correct examples, and a hint facility, e.g., an option to see the beginning of the correct answer.

The trial suggests that the system has the potential to be a useful tool for language study. There was a big difference between the students in how well the system correctly interpreted their intended utterances. Those who were less successful with the ASR expressed frustration, though they were very pleased if they eventually succeeded without resorting to correction by typing. Some were able to adjust their way of speaking to gain distinct improvement, to their great satisfaction. Some special features of Japanese speakers' pronunciation of English were evident. For instance, names containing the "i" sound, "Bill" and "Linda", caused difficulty with ASR, whereas "Monica" was recognized almost 100% of the time. "Would you..." caused great difficulty, whereas "Can you..." went much more smoothly.

6. Discussion

Even from the participants in the pilot study who had less success with ASR, the reaction was generally positive, so the system seems to have potential as a useful tool for the language teacher. It is being made available in the hope that it will be used in both classroom and self-study environments to the benefit of both teachers and learners.

Though in its essence, the system is a facilitator of pattern practice, it is hoped that it can be regarded as a highly enriched version of same, with elements of genuine communicative practice. Interestingly, its two separate parts can be seen as embodying two important concepts in Second Language Acquisition Theory: q-answer provides the user with *comprehensible input* (Krashen, 1981), and q-ask requires the user to produce *comprehensible output* (Swain, 1985). While the learner's need for input is so obvious that nobody would question its central importance in SLA, the importance of producing output has not been universally accepted, with Krashen and others continuing to downplay its significance because of its strong element of conscious learning (Krashen, 1998; Ponniah & Krashen, 2008; Ponniah, 2010; Jarvis & Krashen, 2014). However, Swain has maintained her view, and been supported by others (Swain & Lapkin, 1995; Iwanaka & Takatsuka, 2007; Mennim, 2007; Donesch-Jezo, 2011; Berkner, 2016), so that her output hypothesis remains a seriously-considered concept among the many which have been proposed in SLA theory.

Although the system, with its emphasis on explicit practice focused on grammatical structures, may not appeal to supporters of the *natural approach* (Krashen & Terrell, 1983), it at least provides one of the recommended features of that approach, a stress-free learning environment. This is because it is designed with self-access in mind, though it is also suitable for classroom use. Also, there is much evidence for the usefulness of explicit practice in language learning (DeKeyser, 2010; Lyster & Sato, 2013; Jones, 2018).

The system could probably be classified as dialogue-based CALL, as defined by Bibauw, François & Desmet (2015), though its very limited scope makes it difficult to fit comfortably into any of their four categories. Indeed, many more sophisticated and capable systems have been developed (Xu & Seneff, 2009; van Doremalen, Boves, Colpaert, Cucchiari, & Strik, 2016; Bodnar, Cucchiari, Penning de Vries, Strik, & van Hout, 2017; Sydorenko, Smits, Evanini, & Ramanarayanan, 2018). However, a distinguishing point of the current system is that it is lightweight and can be freely obtained and deployed by anybody who has upload access to a website. They can use it as is, or edit the vocabulary items to their liking. It can then function as one resource among many, in what Gimeno-Sanz (2016) calls “atomized CALL” in contrast to Warschauer’s (1996) “integrative CALL”.

7. Download

The two parts of the system can be tested online and downloaded as zip files from <https://www.mylesobrien.com/q-answer/> and <https://www.mylesobrien.com/q-ask/>. Each zip file includes the HTML file and images folder, which contains all the image files used in the user interface. These may be left unchanged, though users are free to make their own customizations by editing the HTML file or replacing any image files with their own. The vocabulary text files used in the online examples are also included. The teacher will need to edit these, or make new ones, to make persistent adjustments to the vocabulary. The author also gives permission for users to make modified versions of the system for educational use by editing the HTML or JavaScript, provided they do not claim the original or modified system as their own work.

References

- Ashwell, T. & Elam, J.R. (2017). How accurately can the Google Web Speech API recognize and transcribe Japanese L2 English learners' oral production? *JALT CALL Journal*, 13(1), 59-76. Retrieved from <https://journal.jaltcall.org/storage/articles/JALTCALL%2013-1-59.pdf>
- Berkner, V. (2016). Revisiting Input and Output Hypotheses in Second Language Learning. *Asian Education Studies*, 1(1), 19-22. Retrieved from <https://pdfs.semanticscholar.org/c06d/ee67d68c61bbe79b57a6d8e1062a185a3bd3.pdf>
- Bibauw, S., François, T., & Desmet, P. (2015). Dialogue-based CALL: an overview of existing research. In F. Helm, L. Bradley, M. Guarda, & S. Thouësny (Eds), *Critical CALL – Proceedings of the 2015 EUROCALL Conference*, Padova, Italy (pp. 57-64). Dublin: Research-publishing.net. Retrieved from <http://dx.doi.org/10.14705/rpnet.2015.000310>
- Bodnar, S., Cucchiari, C., Penning de Vries, B., Strik, H. & van Hout, R. (2017). Learner affect in computerised L2 oral grammar practice with corrective feedback, *Computer Assisted Language Learning*, 30(3-4), 223-246.
- Daniels, P., & Iwago, K. (2017). The suitability of cloud-based speech recognition engines for language learning. *JALT CALL Journal*, 13(3), 211–221. Retrieved from <https://journal.jaltcall.org/articles/220>
- DeKeyser, R. (2010). Practice for Second Language Learning: Don't Throw out the Baby with the Bathwater. *International Journal of English Studies*, 10 (1), 2010, 155-165. Retrieved from <https://digitum.um.es/xmlui/bitstream/10201/23558/1/114021-452651-1-PB.pdf>

- Donesch-Jezo, E. (2011). The role of output and feedback in second language acquisition – a classroom-based study of grammar acquisition by adult English language learners. *Journal of Estonian and Finno-Ugric Linguistics*, 2(2), 9-28. Retrieved from <http://ojs.utlib.ee/index.php/jeful/article/view/15323/10301>
- Gimeno-Sanz, A. (2016). Moving a step further from “integrative CALL”. What’s to come? *Computer Assisted Language Learning*, 29:6, 1102-1115.
- Iwanaka, T. & Takatsuka, S. (2007). Roles of Output and Noticing in SLA: Does Exposure to Relevant Input Immediately After Output Promote Vocabulary Learning? *Annual Review of English Language Education in Japan*, 18, 121-130. Retrieved from https://www.jstage.jst.go.jp/article/arele/18/0/18_KJ00007108492/pdf-char/en
- Jarvis, H. & Krashen, S. (2014). Is CALL Obsolete? Language Acquisition and Language Learning Revisited in a Digital Age. *TESL-EJ*, 17(4), 1-6. Retrieved from <http://tesl-ej.org/pdf/ej68/a1.pdf>
- Jones, C. (Ed.) (2018). *Practice in Second Language Learning*. Cambridge: Cambridge University press.
- Këpuska, V., Bohouta, G. (2017). Comparing Speech Recognition Systems (Microsoft API, Google API and CMU Sphinx). *Int. Journal of Engineering Research and Application*, 7(3), 20-24. Retrieved from <https://pdfs.semanticscholar.org/2e7e/bdd353c1de9e47fdd1cf0fce61bd33d87103.pdf>
- Krashen, S.D. (1982). *Principles and Practice in Second Language Acquisition*. London: Pergamon. Retrieved from http://www.sdkrashen.com/content/books/principles_and_practice.pdf
- Krashen, S.D. & Terrell, T.D. (1983). *The Natural Approach: Language Acquisition in the Classroom*. San Francisco: Alemany Press. Retrieved from http://sdkrashen.com/content/books/the_natural_approach.pdf
- Krashen, S. (1998). Comprehensible Output. *System*, 26, 175-182. Retrieved from http://www.sdkrashen.com/content/articles/comprehensible_output.pdf
- Lyster, R., & Sato, M. (2013). Skill Acquisition Theory and the role of practice in L2 development. In M.G. Mayo, J. Gutierrez-Mangado, & M.M. Adrian (Eds.), *Contemporary Approaches to Second Language Acquisition* (pp. 71–92). Amsterdam: John Benjamins.
- Mennim, P. (2007). Long-term effects of noticing on oral output. *Language Teaching Research*, 11(3), 265-280.
- Ponniah, R. J. and Krashen, S. (2008). The Expanded Output Hypothesis. *The International Journal of Foreign Language Teaching*, 4(2), 2-3. Retrieved from https://www.academia.edu/738658/The_expanded_output_hypothesis
- Ponniah, R. J. (2010). Insights into Second Language Acquisition Theory and Different Approaches to Language Teaching. *Journal on Educational Psychology*, 3(4), 14-18. Retrieved from <https://files.eric.ed.gov/fulltext/EJ1102366.pdf>
- O'Brien, M. (1997). A computer program to provide practice in questions and answers for learners of English. *Computer Assisted Language Learning*, 10(3), 299-305.
- O'Brien, M. (2017). A freely-available authoring system for browser-based CALL with speech recognition. *EUROCALL Review*, 25(1), 16-25. Retrieved from <https://polipapers.upv.es/index.php/eurocall/article/view/6830>

Swain, M. (1985). Communicative competence: some roles of comprehensible input and comprehensible output in its development. In S.M. Gass and C.G. Madden (Eds.) *Input in second language acquisition* (pp. 235–253). Rowley, MA: Newbury House.

Swain, M. & Lapkin, S. (1995). Problems in Output and the Cognitive Processes they Generate: A Step Towards Second Language Learning. *Applied Linguistics*, 16(3), 371-91. Retrieved from <http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.888.9175&rep=rep1&type=pdf>

Sydorenko, T., Smits, T., Evanini, K., & Ramanarayanan, K. (2018). Simulated speaking environments for language learning: insights from three cases. *Computer Assisted Language Learning*, 32(1-2), 17-48.

van Doremalen, J., Boves, L., Colpaert, J., Cucchiari, C., & Strik, H. (2016). Evaluating automatic speech recognition-based language learning systems: A case study. *Computer Assisted Language Learning*, 29(4), 833-851.

Warschauer, M. (1996). Computer Assisted Language Learning: An Introduction. In S. Fotos (Ed.), *Multimedia language teaching* (pp. 3-20). Tokyo: Logos International. Retrieved from <http://www.ict4lt.org/en/warschauer.htm>

Xu, Y. & Seneff, S. (2009). Speech-Based Interactive Games for Language Learning: Reading, Translation, and Question-Answering. *Computational Linguistics and Chinese Language Processing*, 14(2), 133-160. Retrieved from <https://www.semanticscholar.org/paper/Speech-Based-Interactive-Games-for-Language-and-Xu-Seneff/29c0398871d6b34e9a220e134f90f5cec6dfb86a>

Impact and reactions to a blended MA course on Language Education and Technology

George S. Ypsilandis
Aristotle University, Greece

ypsi@itl.auth.gr

Abstract

The subject of computers in language learning was not covered at a postgraduate level in Greece independently but as an add-on module in more broad programmes, such as *applied linguistics* or *TEFL*. In such programmes, this module was merely scratching the surface of the subject, leaving students with the impression that there was no more to it than learning to run a software programme or an application. The MA programme on *Language Education and Technology* (LET) was the first in the country that aimed to offer a specialised course with all its modules directly related to the area. Furthermore, the programme attempted to incorporate a number of novelties in the personnel involved (experts from six different countries), methods of teaching (blended, through face-to-face, and synchronous web teleconferencing), transparency (as to the use and allocation of the fees and student selection), systems of examination, modes of collaboration, and modules and seminars offered, all directly linked to its title.

The study described here aimed to shed light and estimate the impact of the course on the professional life of its participants through several open and closed questions included in a questionnaire, constructed to register student status before and after the programme, and their opinions on several other programme features. Students scored very positively a) module development, b) the instructors that were involved, c) the modules offered, and d) the knowledge they gained. Some of the students presented their final papers at international conferences, four were accepted in PhD studies in Spain, the UK and Austria, with scholarships from the host institution, while others increased their salaries, or found a new better paid job.

Keywords: Computer-assisted language learning, master's degree, teaching methods.

1. Introduction

A self-supporting programme is difficult to construct especially when it deals with a very specific area of language education. This is because most departments or faculties do not have the commodity of having a sufficient number of scholars working in the same scientific area so as to equip such a programme at an MA level. This explains the above claimed policy of most departments to offer MA programmes in very broad areas of study, which in some cases, resemble a BA degree instead of an MA. A solution to this problem would be to invite experts from other universities, a policy that creates other barriers related to transportation, educational environment and culture, common language, and even administrative issues. This is where technologies become important, especially those which assist synchronous communication and support the creation of hybrid classes and offer assistance to the problem of transportation and educational environment. Big Blue Button (BBB), which is a web conferencing system, was the tool selected due to the ease with which it can engage remote students in online learning by providing real-time sharing of audio files, videos, presentation slides, chat, and desktop with students. The system includes a whiteboard tool that automatically displays annotations back to the students in real-time, while teachers can zoom, highlight, draw and write on presentations to clarify their points. Additionally, there is no limit to the number of webcams that can be connected, and students were able to send and receive instant messages to and from other students or the

instructor, thus being able to ask questions orally or in writing. Lectures could be recorded and be made available for later review. Figure 1 illustrates a typical BBB screen.

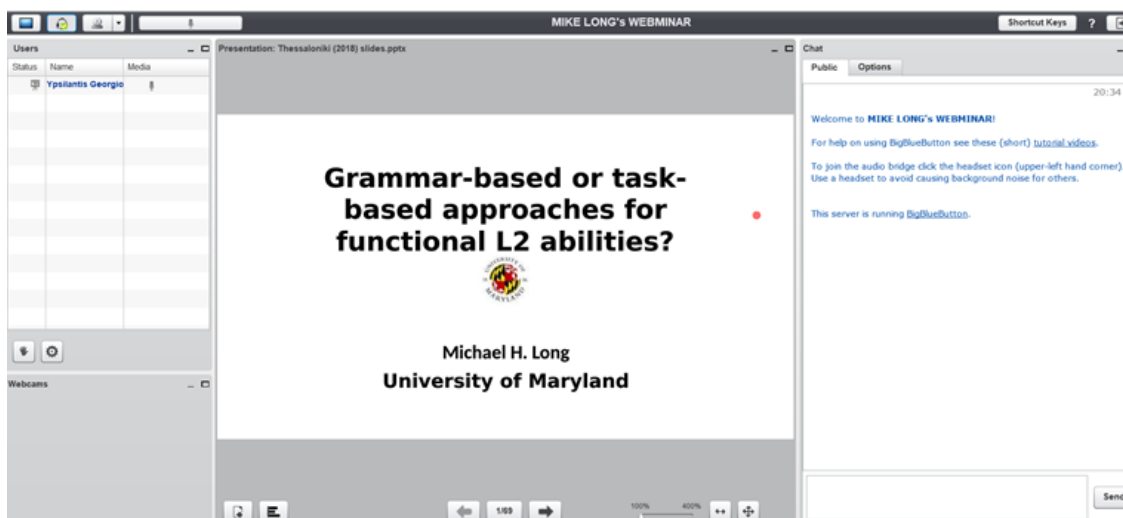


Figure 1. Sample BigBlueButton screen.

As we can see, the screen is divided into three sections. On the left-hand side, the names of the participants appear under the instructor's name; the middle section displays the presentation slides, and any written messages sent to the instructor appear on the right-hand side.

Given the two major problems discussed above, i.e., lack of highly specialised teaching staff and an appropriate educational setting, it was decided that the most suitable solution would be to adopt *blended learning*, allowing “a convergence between face-to-face and technology-mediated learning environments” (Naaj, Nachouki and Ankit, 2012). Thus, instructors from within the university or the country could offer their sessions (mainly practical) in a traditional classroom context, while the international lecturers offered a combination of BBB sessions and intensive face-to-face sessions during a one week visit to the host university, i.e., a form of a blended learning (Badawi, 2009; Naaj, Nachouki and Ankit, 2012). Students were asked to study some of the materials autonomously, through printed or electronic means, as suggested by their instructors.

This study presented here evaluates the master's degree programme on *Technology in Language Education*, by reporting any possible impact of the course on the students' professional life, by requesting information about the participants' previous and current (at the time this paper was written) employment status, as well as their personal opinions regarding how well the programme modules and extra-curricular seminars related to their personal targets, appropriateness of the experts with regard to the modules taught, and the degree of assistance from administration. Finally, students were invited to offer their personal views as to how they felt the programme could be enriched. The outcomes of the study would be of value to both MA course organisers and students alike, as indirectly these would provide a set of criteria to use for MA course design and implementation.

The following sections include an overview of what is understood by *blended learning* and a detailed description of the MA programme, i.e., details of module titles and deployment methods, student selection procedures, etc. Data collected, which is statistically analysed, corresponds to participants who completed the course. The findings are reported and discussed in the last section.

2. The blended learning umbrella

Blended learning (BL) is considered to be an attempt to overcome the weaknesses of face-to-face and distance education (Osguthorpe & Graham, 2003) by offering the spontaneity of face-to-face education, waive the isolation and reduced motivation of distance education reported by Islam (2002), and expand the educational options to learners and teachers alike. The term BL is, to a certain extent, misleading as it suggests that the focus is on learning while most scholars (Garnham & Kaleta, 2002; Li & Zhao, 2004; Bonk & Graham, 2006; Chan, 2008) define it as a method combining different teaching methods and approaches (i.e., traditional teaching techniques), with the assistance of information technology (Gulbahar & Madran, 2009; Picciano, 2006) under a new pedagogical approach. BL is therefore proclaimed to have come about to support a traditional teaching paradigm, the effect of which has arguably increased. In this respect, we could conclude that BL focuses on teaching more than learning although both aspects of education are assumed to progress in harmony, with the latter (learning) being an outcome of the former (teaching). The above claim may also be seen in Delialioglu and Yildirim (2007), who state that BL focusses on the systematic usage of and strategic engagement with electronic devices to achieve *teaching* targets and individual *learning* goals of the students. Li and Zhao (2004), and Garrison and Kanuka (2004) also maintain that effectiveness of traditional *teaching* is increased in blended *learning* environments. This line of thought has been strongly debated in Ypsilandis (2005, 2006), and Oliver and Trigwell (2005) where, after a detailed contrastive discussion of the concepts of *teaching* and *learning*, Ypsilandis concludes that teaching does not always result in learning, while learning can be achieved without any formal blended instruction. Indeed, Oliver and Trigwell (2005) recommend that the term 'learning' in BL should be abandoned as "...learning from the perspective of the learner, is rarely, if ever, the subject of blended learning. What is actually being addressed are, forms of instruction, teaching, or at best, pedagogies." It is also suggested that most of the attention in BL is on the appropriate usage and application of technology, which evolves rapidly, and offers many tools, originally designed for educational purposes or for other commitments that can be also be pedagogically exploited. BL has also been thought of as an in-between stage, which was not going to last long, en-route to fully-fledged computer-based instruction, as technology was progressing fast. However, despite the rapid development of computerised technology and the vast number of applications constantly emerging, both dedicated and non-dedicated, human involvement is still present in blended learning, which involves traditional teaching, in the form of lectures, seminars or workshops.

The theoretical background supporting BL comes from, and inevitably links to, learner autonomy, learner needs and tailored learning in which the individual learner's needs are met through increased and self-paced participation (Thorne 2003). Self-motivation, self-management and self-regulated learning are also thought to be improved (So & Brush, 2008), while different pedagogical models (constructivism, behaviourism and cognitivism) may underpin a BL educational setting (Driscoll, 2002). Finally, Singh (2001) proclaims a thoughtful and careful use of the 'right' learning technologies in association with the 'right' skills, time and personal learning style. In conclusion, BL cannot be seen as a methodology but rather as a teaching approach that attempts to optimise the use of technology to support both autonomous learning and teacher-led instruction, aiming to increase the amount and/or the quality of learning.

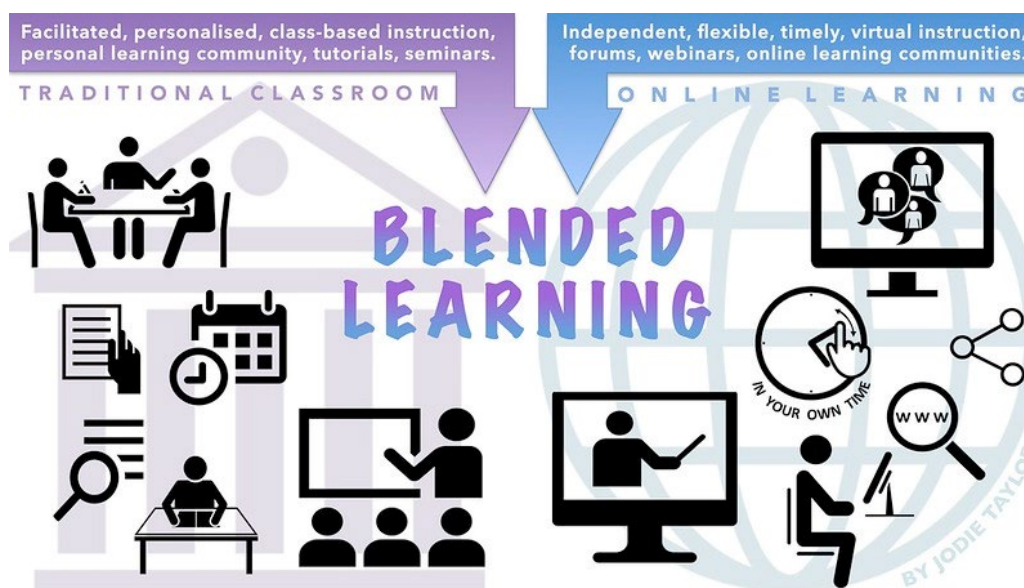


Figure 2. Blended learning (Attribution: <https://www.flickr.com/photos/jodieinblack/29155993523>).

The MA programme on *Language Education and Technology* attempted to bring new potential to the concept of BL by combining the use of technological tools with the expertise of scholars from different cultural backgrounds and academic institutions who were willing to add a global perspective to its scope. The term *blended*, in this case, takes the form of not only a mix of traditional and electronic tools but also a blend of experts that contributed to the course without having to live in the location of the institution, the Aristotle University of Thessaloniki in Greece. Consequently, it was essential that the course be offered through an alternative mode of delivery, such as online or hybrid instruction, alternatively scheduled classes, in alternative locations (e.g. on and off-campus, face-to-face or virtual) to accommodate the instructors and the students, some of whom were in employment during their studies.

3. The MA programme

In 2015, government regulations in Greece permitted state universities to prepare and offer self-supporting postgraduate programmes, at MA level, to serve full-time fee-paying registered students. This change of position was welcomed by most academics and students alike though it was not seen positively by left wing political parties, and their supporters within academic institutions, whose views were that education should be offered free of charge to all students. These new regulations offered the possibility for this programme to materialise based on two objectives: a) to facilitate Greek students' access to renowned instructors in the fields of applied linguistics and computer-assisted language learning (CALL), making use of the University's existing technological infrastructure, and b) put together a syllabus comprising cutting-edge technology and related pedagogies. Several scholars from around the world were invited and accepted the challenge. The MA programme was launched in October 2015 and the students who were selected were predominantly teaching professionals (from the public and private sector) who were either full-time employees, mid-career professionals or students with specialized goals who wished to pursue an academic research career. There was also a very small number of applicants from other backgrounds, i.e., theatrical studies, accounting, and education. Student selection was achieved through a set of criteria (i.e., BA grade, years of teaching experience, experience in academic writing, etc.) presented on an Excel file grid, which was then uploaded onto the Departmental intranet, with the applicants in a norm referenced mode to ensure transparency of decisions and fair treatment.

The programme aimed to cover two basic aspects related to language education and technology; the first was an introduction to academic research, for students wishing to pursue an academic career, and the second was the practical side, for those working as language teachers or software

developers. The modules selected followed the above rationale and were distributed equally in two semesters (the taught part), while a third semester was devoted to writing the final MA dissertation (autonomous supervised research). The modules covered the following areas: experimental research methods and statistics, second language acquisition theories, mobile-assisted language learning, the internet and language education, internet technology, language teaching theories and CALL, learner autonomy, instructive vs incidental learning, databases and language applications, and massive open online language courses (Language MOOCs).

NEW STARTING FROM OCTOBER 2015!

**MASTER'S COURSE IN
LANGUAGE EDUCATION
AND TECHNOLOGY**

MODULES	ECTS
Winter semester	
Research Methods and Statistics	8
An Introduction to CALL and the History of CALL	8
The Internet and Language Education	8
SEMINAR I	6
Mobile-Assisted Language Learning	
Summer semester	
Databases and Language Applications	8
Internet Technology	8
Second Language Acquisition Theories	8
SEMINAR II	6
Massive Open Online Courses	



LIST OF INSTRUCTORS	GUEST SPEAKERS
Panos Arvanitis (Greece)	David Little (Ireland)
Elena Bárcena (Spain)	Mike Long (USA)
Ana Gimeno-Sanz (Spain)	Heiner Boettger (Germany)
Mirjam Hauck (UK)	
Anna Mouti (Greece)	
Panos Panagiotidis (Greece)	
Aggeliki Psaltou-Joycey (Greece)	
Glen Stockwell (Japan)	
George S. Ypsilandis (Greece)	

**ARISTOTLE UNIVERSITY OF THESSALONIKI
GREECE**

OCTOBER 2015 – JUNE 2016

FEES: 1500 EUROS (in three installments)

INFO: George S. Ypsilandis ypsi@itl.auth.gr

Figure 3. MA course flyer the year it was launched.

The entire programme offered 90 ECTS from the above modules and seminars. Additional seminars were also offered on topics related to language teaching methodologies with invited speakers. These seminars did not offer extra ECTS; however, students received a certificate of attendance. Finally, a practicum was arranged for interested parties at the ATLAS Research Group in the Departamento de Filologías Extranjeras y sus Lingüísticas, at the Faculty of Philology of the National Distance Education University based in Madrid (Spain), also known as UNED, through the ERASMUS mobility programme, under the supervision of Prof. Elena Bárcena. The instructors selected were some of the most renowned and highly acclaimed international scholars and researchers in the field of Computer-Assisted Language Learning (CALL): Ana Gimeno-Sanz (Professor of English and Applied Linguistics at the Universitat Politècnica de València, Spain, and current President of WorldCALL, who also acted as co-coordinator of the programme), Mirjam Hauck (Senior Lecturer in the Department of Languages at the Open University, UK, and current President of EuroCALL), Prof. Glen Stockwell (Associate Dean of the Faculty of Law at Waseda University, Japan), Prof. Elena Bárcena (Professor of English at UNED, Spain) and five Greek nationals: Thomas Vougiouklis (Professor Emeritus at the Democritus University of Thrace), and from the Aristotle University of Thessaloniki: Agneliki Psaltou-Joycey (Professor Emeritus), George Ypsilandis (Professor and programme coordinator), Panos Arvanitis and Panos Panagiotidis (Associate professors). Additional seminars, by well-known scholars who were invited to teach on campus or through webinars, were offered. Those who participated were Mike Long (USA), David Little (Ireland), Phil Hubbard (USA), Heiner Boettger (Germany), and Cornelia Illie (Sweden). The working language of the course was English.

The content of the various modules was uploaded onto the host university's learning management system, which is based on Moodle.

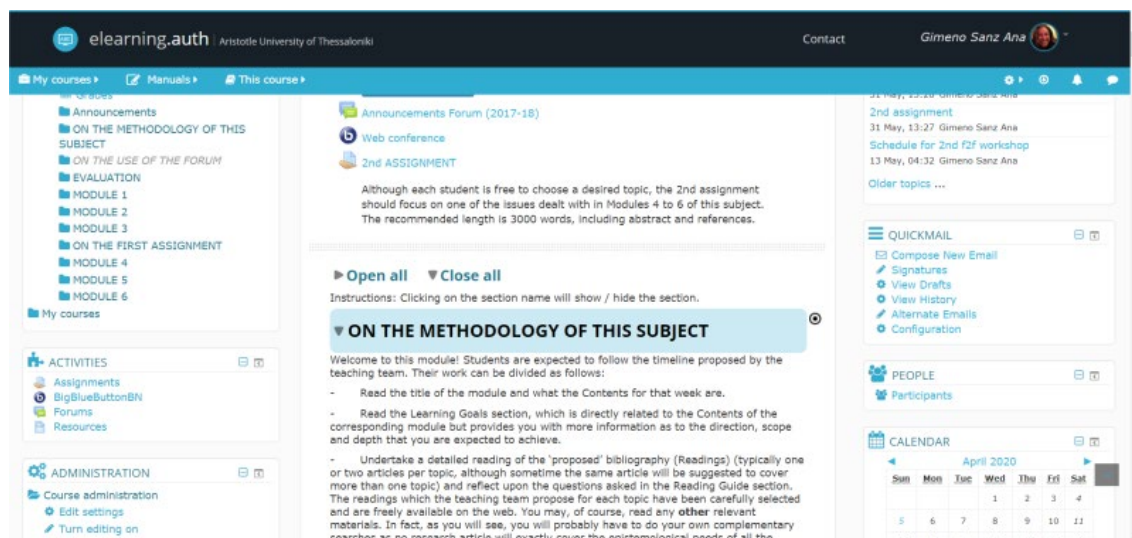


Figure 4. Moodle platform used for the course content.

In a 13-week semester, each instructor offered 39 hours of teaching based on one of the following or a combination of these: 1) through BBB used from home, 2) by traditional face-to-face class meetings, and 3) through autonomous study (up to four weeks per semester). Autonomous study weeks were used by the instructors to engage students in studying a specific topic independently and present it in class. Lastly, external instructors were scheduled to visit the host university for a week and offer a series of face-to-face seminars or workshops related to the module they were teaching. Information, such as *module description*, the *teaching targets* and *learning outcomes*, the *evaluation methods* and the *marking system* together with a *detailed plan of meetings*, *autonomous study weeks*, and *planned visits* for each module was presented to the students at the beginning of each semester. Programme modules were organised in one of, or in a combination of, the following three formats:

1) The serial development was a typical flow from A to B, to C, to D, etc. This was the type of development that was used by most instructors.

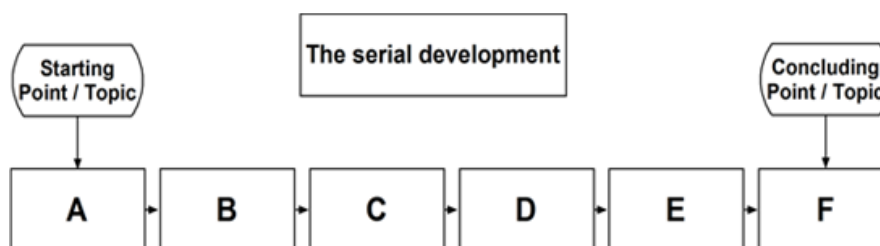


Figure 5. Serial development model.

In this module deployment, the starting and concluding points differ and students are expected to read or follow the material in a serial fashion, i.e., one was not advised to read material in D before having read materials in A, B, and C.

2) In a star development, where the main topic is in the centre and the relevant topics (that can be independent of each other) are presented in the form of a star. Students can therefore select a topic without necessarily having to read the other ones.

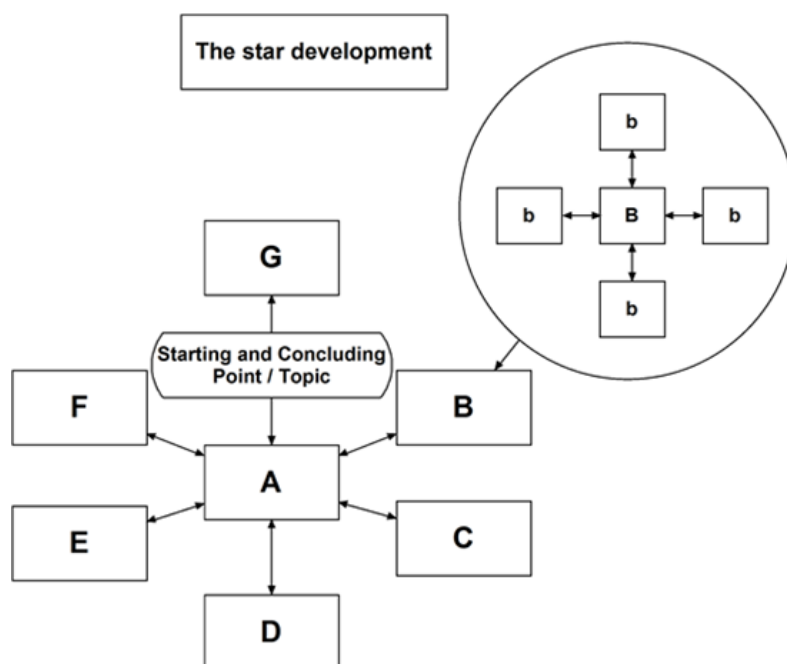


Figure 6. Star development model.

Here, there is no starting or ending point. The main topic lies in the centre and all the other topics are spread around and contribute to its understanding. Unlike the serial development, there was no need for the students to cover all the other related topics if they felt they were already equipped with the required knowledge. All the topics could be covered independently and one could start from any one a student wished. Furthermore, a topic (e.g. B) could include material developed in another star or even serial or cyclical developments.

3) The cyclical development was recommended for topics in which one could start from one point or topic and return to it with the added knowledge acquired in the circle.

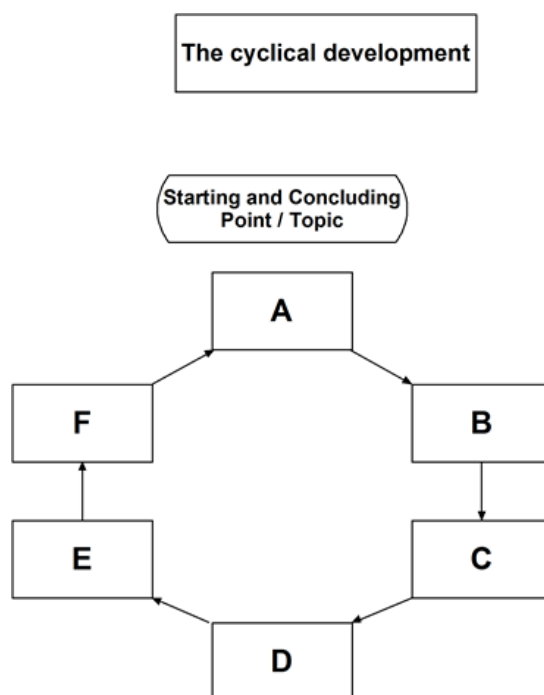


Figure 7. Cyclical development model.

Although this may resemble a serial development, this type of development differs in that the starting and concluding points are the same and the students arrive there with the added knowledge of the subject, e.g., if one were to teach the use of *Hot Potatoes* one may start from a textbook (A), spot exercises and drills in it (B) think what specific POTATO to use (C), make the necessary adaptations to fit in the POTATO (D), input the material in the POTATO selected (E), run it to see if it works properly and make corrections (F), look at the final product and integrate the material in his/her teaching in relation to the textbook in a relevant manner (A). All intervals in the circle may include materials to develop the necessary skills to complete them.

4. Method

4.1. Participants

Fifty students, who were registered in the first two years of the course, were initially targeted; 6 of those had not completed the programme at the time of the study and were excluded from the sample. A final sample of 44 subjects were approached and 20 of those responded to the appeal (return rate of 45.4%).

4.2. Design and procedure

The study proceeded with the administration of a questionnaire to students of the first two years who had completed the programme. Participants were approached by email and were asked to fill in the survey, designed for the purpose of the study, with a student acting as a mediator to collect responses and thus ensure anonymity.

4.3. Instruments and materials

The instrument used was a questionnaire with 13 questions: 8 open and 5 based on a ratio scale –*ravdos*– (Kambakis-Vougiouklis & Vougiouklis, 2008; Kambakis-Vougiouklis, Nikolaidou, & Vougiouklis, 2017). The *ravdos* scale was chosen instead of the traditional Likert scale because the former offers a very precise continuum for measuring personal perceptions that takes into account not only integers but also decimals. The SPSS (v.25) statistical package was employed for data analysis.

5. Analysis

The first part of the analysis offers some general data concerning student participation in various conferences in Europe, supported with a 300 euro scholarship from the programme. The rest of the analysis follows the questions as they appeared in the survey. In particular, frequencies of the variables and correlations between them are offered, supported with graphs where deemed necessary. The percentages presented in the relevant tables are those without the missing items which were treated pairwise.

5.1. General data

Looking at the records kept by the financial manager, it was found that a total of 15 students (34%), out of the 44 who completed the course, presented their research findings in conferences in Europe (Belgium, France, Bosnia and Herzogavina, Spain and Austria). This detail supports the impact of the course on the students' academic life, described in another study by Ypsilandis (2018).

5.2. Frequencies and correlations between previous and current employment

Reporting on employment prior to and after taking the MA course allows us to assess the effect of the study programme on the subjects' professional career (Table 1).

Table 1. Students' pre- and post-course employment.

1		P. J.	C. J.
2	Unemployed	5	1
3	Language Teacher	11	12
4	Account assistant	1	0
5	Librarian	1	1
6	Researcher	1	1
7	Primary School Teacher	1	1
8	Human Resource Trainer	0	1
9	Student	0	3
10	Total	20	20

P.J. = Previous Job / C.J.= Current Job

As expected, most participants (11, 55%) were language teachers working in either the private or the public sector. Five (25%) were unemployed, while there was an accountant assistant, a librarian, a researcher, and a primary school teacher. It becomes clear (Table 1) that only one participant remained unemployed after completing the MA. The Population Pyramid (1) below, presents changes in subjects' employment before and after the programme schematically and in more detail.

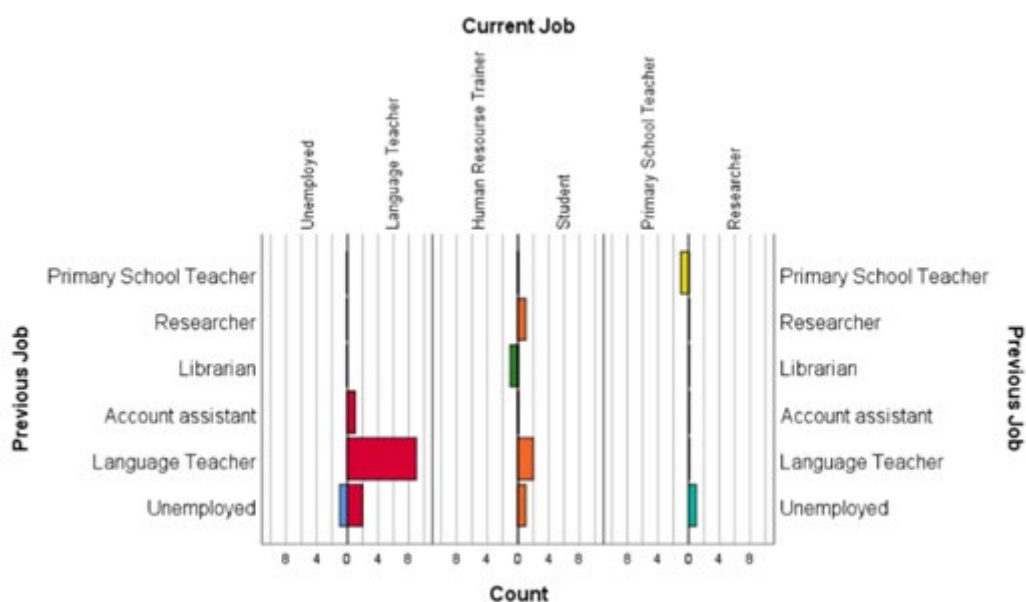


Figure 8. Population pyramid indicating previous and current jobs.

Prior employment is registered on the left-hand side of the vertical lines while post programme occupation is shown on the right-hand side. Note that, the accountant assistant found a teaching position, and the same was true for 2 of the previously unemployed. Two of the participants previously registered as language teachers continued on a PhD programme and the same was found for 1 previously unemployed person. In addition, one changed profession, as she found a

job as a human-resource trainer and one subject remained a language teacher though in a much more promising and better paid position as a language teacher in China (and also pursued a PhD). The others remained in their previous professions. In Figure 8, we can see that students mainly aspired to become a language teacher or pursue a PhD.

A cross-tabulation test was used to explore the correlation between the two nominal variables, i.e., previous and current employment (Table, 2).

Table 2. Correlation between previous and current employment.

Chi-Square Tests						
	Value	df	Asymptotic Significance (2-sided)	Exact Sig. (2-sided)	Exact Sig. (1-sided)	Point Probability
Pearson Chi-Square	51.091 ^a	25	.002	.027		
Likelihood Ratio	25.348	25	.443	.038		
Fisher's Exact Test	41.648			.034		
Linear-by-Linear Association	1.967 ^b	1	.161	.165	.103	.014
N of Valid Cases	20					

a. 35 cells (97.2%) have expected count less than 5. The minimum expected count is .05.

b. The standardized statistic is 1.403.

A statistically significant correlation between the two variables was detected (Fisher's Exact Test = .03, $p < .05$, Chi Square = 51, DF=25). Cramer's V ($V = .71$) statistic, which shows the strength of the relationship between the tested variables, shows that this was particularly strong, at the level of *probably measuring the same concept*. This means that, those who were already in the teaching profession stayed in it after the MA programme with three more persons becoming language teachers. The results for a related closed question as to whether they found the programme had added to their professional development may be further enlightening (Table 3).

Table 3. Question enquiring whether the MA course had had an effect on the students' professional development.

Has MA added to Professional Development					
		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	No	3	15.0	15.0	15.0
	Yes	17	85.0	85.0	100.0
	Total	20	100.0	100.0	

Seventeen subjects (85%) selected 'yes' while 3 (15%) answered a categorical 'no' to this question. A further open question, which explored the opinions of the subjects asking them to declare the exact gained knowledge or skills they had acquired after completing the programme, is presented in Table 4.

Table 4. Skills or knowledge acquired after taking the MA course.

		How did it Help?			
		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	Innovative Approaches	5	25,0	31,3	31,3
	Apply Research	3	15,0	18,8	50,0
	Got a job	4	20,0	25,0	75,0
	Extra Degree	2	10,0	12,5	87,5
	Did not help	1	5,0	6,3	93,8
	Got more money	1	5,0	6,3	100,0
	Total	16	80,0	100,0	
Missing	999,00	4	20,0		
Total		20	100,0		

Looking at the frequencies of the answers provided, it is possible to conclude that the opinions were quite spread. Five (31.3%) declared that the programme enriched their professional profile with innovative language teaching approaches, 4 (25%) found a job (not necessarily a teaching position), 3 (18.8%) reported they had learned how to apply research to their profession, 2 (12.5%) declared that they had simply added another qualification to their portfolio, and 1 (6.3%) increased her salary. There was only 1 (6.3%) subject who clearly stated that the programme did not help her professional career while 4 (25%) did not provide an answer to this question.

5.3. Relationship between programme modules and personal targets

A ratio-scored question based on Kambakis-Vougiouklis and Vougiouklis (2008) *ravdos* scale (from 1 to 10) was used to record the subjects' opinions to the question as to whether the course modules were relevant to their personal targets (Table 5). Students reported the following:

Table 5. Correlation between course modules and students' learning target.

Statistics		
Association of Modules with Learning Target		
N	Valid	20
	Missing	0
Mean		8.0000
Median		8.0000
Mode		8.00
Std. Deviation		1.45095

The mean, median (score which splits the sample in half) and the mode (most recorded score) were all found to be very high at 8/10 (N=20). Standard deviation, which quantifies the amount of variation of a set of data values, was found to be at 1.4, which confirms the positive reaction to this question. Figure 9 provides the exact frequencies.

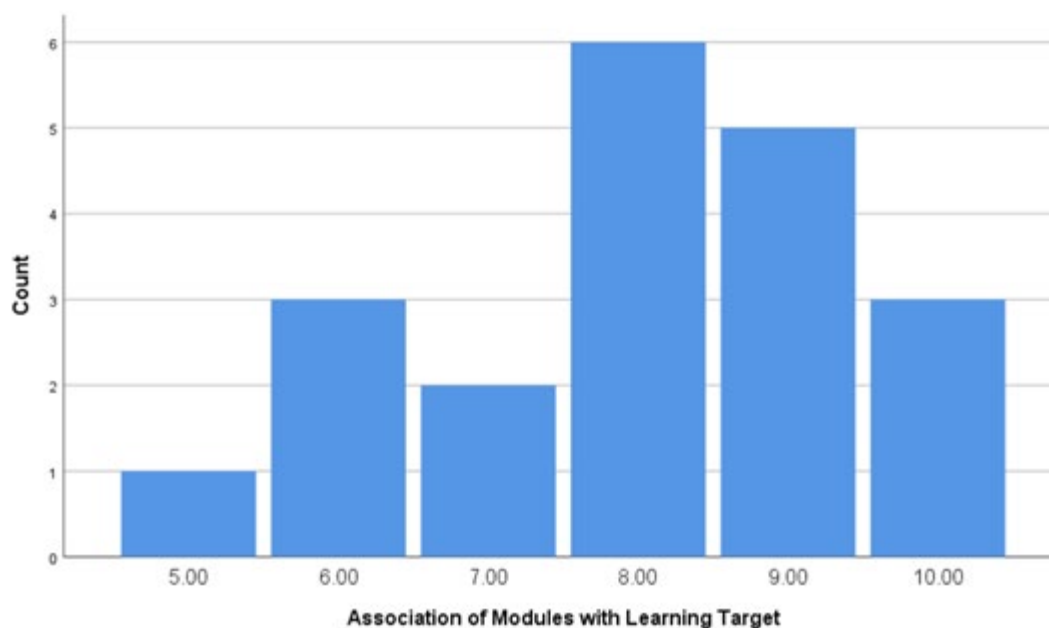


Figure 9. Alignment between the modules and the learning target.

Most of the subjects stated that the programme modules were relevant to their personal targets and selected the highest scores on the ravidos scale (8-9-10). Although all the scores are above the 5-median level, it appears that not all personal targets were covered. In Ypsilandis (2018), the non-covered personal targets of the programme were specified to be: a) increasing quality of learning in their classes, and b) making their classes more interesting.

5.4. Association of extra-curricular seminars to personal targets

As in the previous question, the same ravidos scale was used to register the subjects' views regarding the relevance of the extra-curricular seminars to their personal targets (Table 6).

Table 6. Relevance of the extra-curricular seminars in relation to personal targets.

Statistics		
Association of Seminars to Learning Targets		
N	Valid	20
	Missing	0
Mean		6.9000
Median		7.5000
Mode		8.00
Std. Deviation		2.07491

In this case, the mean is much lower (mean=6.9) than the one in the previous section, closer to the middle of the ravidos scale. Similarly, the median is at 7.5, while mode was found at 8. Standard deviation is at 2 which reflects a wider dispersion of data values. Table 7 offers the details of students' scores.

Table 7. Correlation between the seminars and the students' learning targets.

Association of Seminars to Learning Targets					
		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	2.00	1	5.0	5.0	5.0
	3.00	1	5.0	5.0	10.0
	5.00	3	15.0	15.0	25.0
	6.00	2	10.0	10.0	35.0
	7.00	3	15.0	15.0	50.0
	8.00	6	30.0	30.0	80.0
	9.00	3	15.0	15.0	95.0
	10.00	1	5.0	5.0	100.0
Total		20	100.0	100.0	

Note that the spread of responses is wider than the one in the previous section (as indicated above by a higher standard deviation), despite the fact that most responses are, again, in the higher part of the ravidos scale. There are two particularly negative responses (scoring 2 and 3 on the ravidos scale) and 3 responses in the middle of the scale. This means that the students did not feel that the extra-curricular seminars were as relevant for them as the modules offered within the course. This is depicted in Figure 10.

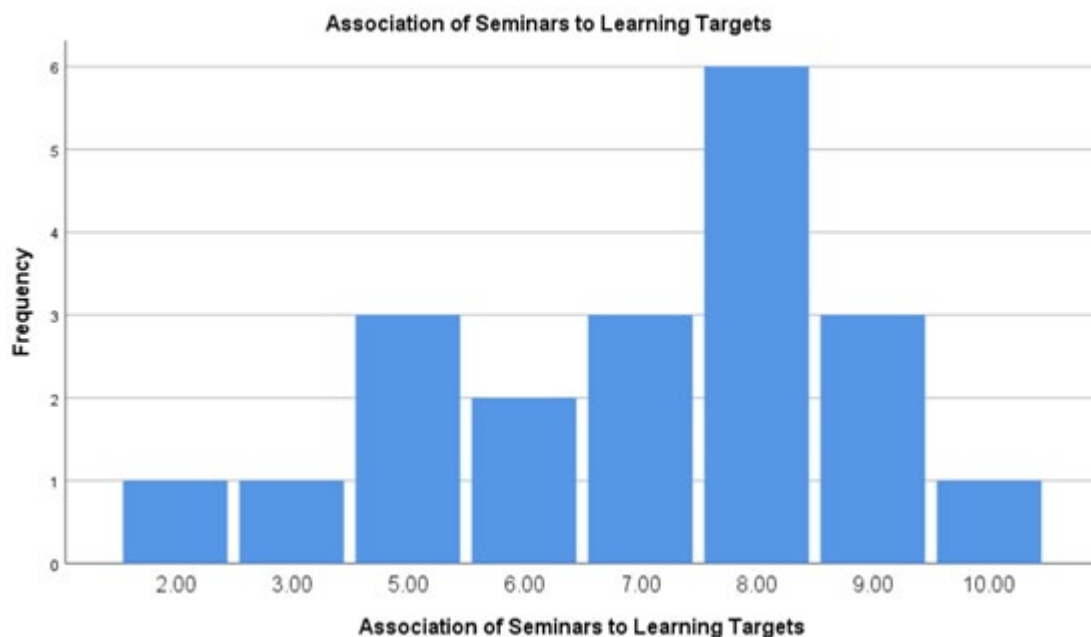


Figure 10. Correlation between the seminars and the students' learning targets.

As we can see, most responses were above the middle section of the scale although the distribution is not symmetrical. There seem to be two concentration points; the first is found at point 8 of the ravidos scale, while the second is around point 5, which confirms earlier claims that the extra-curricular seminars and the modules were not equally relevant for the students.

5.5. Adequacy of the experts and the modules taught

Another question related to the adequacy of the experts selected to deliver the modules offered within the course. This question was also scored on the same ravidos scale (Table 8).

Table 8. Satisfaction with the experts who delivered the modules.

Statistics		
Association of Personnel with Taught Modules		
N	Valid	20
	Missing	0
Mean		9.1000
Median		9.0000
Mode		10.00
Std. Deviation		.85224

Mean, median and mode were all above point 9 of the ravidos scale and thus extremely positive. Standard deviation was at .85, which confirms this positive reaction. Table 9 is more illuminating.

Table 9. Adequacy of the experts in relation to the modules they delivered.

Association of Personnel with Taught Modules					
		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	8.00	6	30.0	30.0	30.0
	9.00	6	30.0	30.0	60.0
	10.00	8	40.0	40.0	100.0
	Total	20	100.0	100.0	

Notice that the majority of the students' votes are at the maximum of the scale and the rest are equally distributed between points 8 and 9 (6 responses each). This is also portrayed in Figure 11.

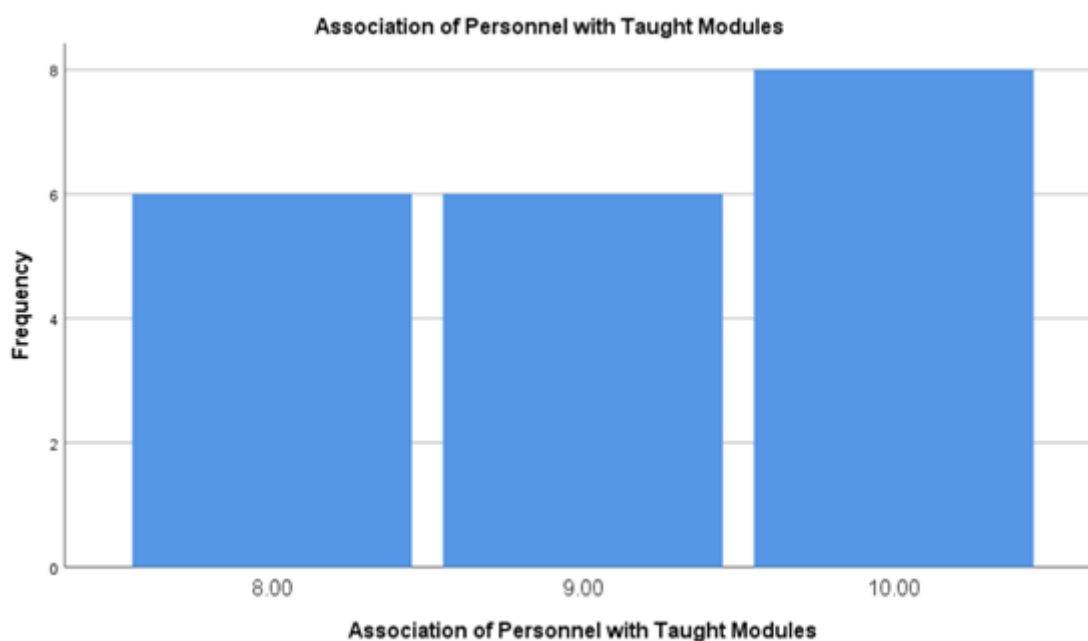


Figure 11. Satisfaction with the experts in relation to the modules they delivered.

Note that there are no responses lower than 8 on the ravidos scale, a reaction that confirms the positive attitude towards the specialists involved in the programme.

5.6. Assistance from administrative staff

The administrative tasks of the programme were shared between two individuals. One who was in charge of departmental administrative tasks, carried out by the secretary of the department running the MA course, and the other in charge of financial issues, who was also responsible for liaising with the students whenever necessary regarding the study programme. Two separate questions registered the students' reactions to administration. The first reflected their opinion in relation to the secretary of the department and the second their views regarding the finance officer (Table 10).

Table 10. Degree of satisfaction with administration.

		Statistics	
		Secretariat Helpful	Financial Office Helpful
N	Valid	20	20
	Missing	0	0
Mean		4.7500	9.5000
Median		5.0000	10.0000
Mode		1.00	10.00
Std. Deviation		2.69258	1.10024

The difference in the students' reactions regarding administration is striking. While secretarial duties were deemed average on the ravidos scale (mean=4.7, mode=1), financial and communication tasks received a very high score (mean=9.5), with the mode value being at 10. Note, also that the standard deviation for the secretarial duties is at 2.6 confirming the higher dispersion of the data while it is at 1.1 for finance and communication, which shows the concentration of the data at the two highest possible points of the ravidos scale.

5.7. Additions to the course

The final open question reflects the students' recommendations for improvement. Table 11 displays the details.

Table 11. Recommended additions to the MA programme.

		What would I add?			
		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	More Seminars	3	15.0	15.8	15.8
	More Practical Courses	5	25.0	26.3	42.1
	More choice of modules	3	15.0	15.8	57.9
	Seminars and Practice	6	30.0	31.6	89.5
	More Practice and Modules	1	5.0	5.3	94.7
	Better Equipment	1	5.0	5.3	100.0
	Total	19	95.0	100.0	
Missing	999.00	1	5.0		
Total		20	100.0		

The majority of the students claimed that they would have liked to have had more seminars and modules devoted to practical topics (6-31.6% and 5-26.3% respectively). Students also pointed out that they would have liked to have had more optional modules to choose from (3, 15.8% for both preferences).

The correlation between the seminars with regard to the learning targets of the programme was not clear to all subjects. Seminars were intended to provide a more in-depth knowledge of the theoretical background related to language education. This was not clear for the students possibly due to lack of information linking the two (educational technologies and language teaching). This finding relates to claims registered in Table 9 where we can see that 23.6% of the respondents clearly stated that they would have wished more modules that are practical in focus based on applications for language teaching. This reaction could be due to the enthusiasm felt by new teachers for practical activities immediately useable in class.

6. Summary and conclusions

The MA course on *Language Education and Technology* delivered at the Aristotle University in Thessaloniki (Greece) was found to have a significant impact on the professional careers of its participants. It either opened job opportunities to those unemployed or increased their professional skills qualitatively with innovative approaches and research skills. Despite the positive claims of module relevance to their personal targets, the subjects were not all equally convinced about the relevance of the extra-curricular seminars to the course. This was possibly due to the lack of information and guidance compared to that received about the course modules, i.e., information about instructor, description of module, learning outcomes, structure, methodology, assessment, and bibliography. Had this information been provided, reactions may have been different and more positive.

Results as to the specialised instructors who participated in the MA were very positive indeed, similar to that reported in Ypsilandis (2018), where the quality of instructors was registered as the main strength of the course and received the highest score of the variables selected to be evaluated. Reactions to this variable reinforce the idea of the benefits of organising a blended learning programme with scholars who can bring their expertise to the programme regardless of their location, rather than only relying on departmental staff, something that can easily be supported by educational technology to make it feasible. Ideas for additions to the programme related to practical modules and seminars to increase the choices from which students could select. This opinion is probably due to the fact that most of the students were in-service language teachers seeking new ideas for their teaching practices.

The fact that the administrative duties conducted by the department scored low supports the idea that, especially in an MA course where many of the enrolees may be simultaneously working and studying, students need to be well-informed of all the administrative procedures and need a clear idea of the bureaucratic requirements.

We can therefore conclude that the programme provided an overall positive experience for its participants, opened new professional directions for them, provided them with opportunities to present at international conferences, undertake a follow-up practicum, or continue conducting research in an international PhD programme.

References

- Badawi, M. F. (2009). Using blended learning for enhanced EFL prospective teachers' pedagogical knowledge and performance. *Proceedings of the Learning & Language Conference – The Spirit of the Age*. Cairo: Ain Shams University.
- Bonk, C.J. & Graham, C.R. (2006). *The handbook of blended learning environments: Global perspectives, local designs*. San Francisco: Jossey-Bass/Pfeiffer.

Chan, C. T. & Koh, Y.Y. (2008). Different Degrees of blending benefit students differently: A Pilot Study. *Proceedings of the EDU-COM 2008 International Conference, 19-21 November 2008*. Retrieved from <http://ro.ecu.edu.au/ceducom/7>.

Delialioglu, O. & Yilkirim, Z. (2007). Student's Perceptions on Effective Dimensions of Interactive Learning in a Blended Learning Environment. *Educational Technology & Society*, 10(2), 133-146.

Driscoll, M. (2002). Blended learning: Let's get beyond the hype. *E-learning*, 1(4), 1-4.

Garnham, C. & Kaleta, R. (2002) Introduction to hybrid courses. *Teaching with Technology Today*, 8(6), 1-2.

Garrison, D.R. & Kanuka, H. (2004). Blended learning: Uncovering its transformative potential in higher education. *The Internet and Higher Education*, 7(2), pp. 95-105.

Gulbahar, Y. & Madran, O. (2009). Communication and collaboration, satisfaction, equity, and autonomy in blended learning environments: A case from Turkey. *International Review of Research in Open and Distance Learning*, 10(2), 1-22.

Islam, K. (2002). Is e-learning floundering: Identifying shortcomings and preparing for success. *E-Learning Magazine*, pp. 22-26

Kambakis-Vougiouklis, P., Vougiouklis, T. (2008). Bar instead of scale. *Ratio Sociologica* , 3, 49-56.

Kambakis-Vougiouklis, P., Nikolaidou, P. & Vougiouklis, T. (2017). Questionnaires in Linguistics Using the Bar and the H v-Structures. In Maturo, A., Hoskova-Mayerova, S., Soitu, D.T. & Kacprzyk, J. (Eds). *Recent Trends in Social Systems: Quantitative Theories and Quantitative Models. Studies in Systems, Decision and Control*, 66. Springer International Publishing AG Switzerland.

Li K. & Zhao J. (2004). The Theory and Applied Model of Blended Learning. *E-Education Research*, 7, 1-6.

Naaj, M., Nachouki, M. & Ankit, A. (2012). Evaluating student satisfaction with blended learning in a gender-segregated environment. *Journal of Information Technology Education: Research*, 11(1), 185-200.

Oliver, M. & K. Trigwell (2005). Can 'Blended Learning' Be Redeemed? *E-Learning*, 2, 17-26.

Osguthorpe, R. & C.R. Graham (2003) . Blended Learning Environments, Definitions and Directions. *The Quarterly Review of Distance Education*, 4(3), pp. 227-233.

Picciano, A.G. (2006). Blended learning: Implications for growth and access. *Journal of Asynchronous Learning Networks*, 10(3), 95-102.

Singh, H., & Reed, C. (2001). *A white paper: Achieving success with blended learning*. Centra software, 1, 1-11.

So, H. J. & Brush, T.A. (2008). Student perceptions of collaborative learning, social presence and satisfaction in a blended learning environment: Relationships and critical factors. *Computers & Education*, 51(1), 318-336.

Thorne, K. (2003). *Blended learning: How to integrate online and traditional learning*. London: Kogan Page.

Ypsilandis, G.S. (2005). Language Teaching, Language Learning: Current Trends and Practices. Keynote published with the proceedings of the 1st ESP Conference on *Teaching English For Specific Purposes: A Trend or A Demand?* Ziti Publications, pp. 31-40.

Ypsilandis, G.S. (2006). On feedback provision strategies in CALL software. Keynote published in the Proceedings of the 5th International Conference on *Motivation in Learning Language for Specific and Academic Purposes*. Thessaloniki: University of Macedonia Press.

Ypsilandis, G.S. (2018). The MA program on Language Education and Technology: A Global Endeavour. *Proceedings of the International Conference on Innovation in Language Learning*. Pixel Publications. Available from <https://conference.pixel-online.net/ICT4LL>.

Appendix

Questionnaire used to collect data.

1. What was your job before you started the MA programme?

2. What is your job now?

3. Do you think the MA programme has assisted you in:

YES, in my job		NOT in my job	
Please, write in what way.			
YES, academically		NOT academically	
Please, write in what way.			

4. Please state 3 strong points of the programme.

1	
2	
3	

5. Before I started the programme, my personal aims were to:

1	
2	
3	

6. When I completed the programme, I learned:

1	
2	
3	

7. Were the modules of the programme related to your personal goals? (Please mark selected cell with an X. Left is negative and right is positive)

8. Were the seminars offered during the programme related to your personal goals? (Please mark selected cell with an X. Left is negative and right is positive)

9. Were the personnel of the programme suitable to teach the modules? (Please mark selected cell with an X. Left is negative and right is positive)

10. In terms of administrative tasks, was the department secretariat helpful? (Please mark selected cell with an X. Left is negative and right is positive)

11. In terms of financial management, was the department financial service helpful? (Please mark selected cell with an X. Left is negative and right is positive)

12. The skills I have acquired during the course are:

1	
2	
3	

13. I would add the following modules to the course:

1	
2	
3	

Websites

BigBlueButton <https://bigbluebutton.org/>

EuroCALL <https://eurocall-languages.org/>

Hot Potatoes <https://hotpot.uvic.ca/>

Moodle <https://moodle.org/?lang=en>

WorldCALL <http://www.worldcall.org>

The Linguacuisine Project: A Cooking-based Language Learning Application

Paul Seedhouse*, Phil Heslop**, Ahmed Kharrufa***, Simin Ren**** and Trang Nguyen*****
Newcastle University, UK

*paul.seedhouse@ncl.ac.uk | **philip.heslop@ncl.ac.uk | ***ahmed.kharrufa@ncl.ac.uk |
****s.ren3@ncl.ac.uk | *****T.T.T.Nguyen3@newcastle.ac.uk

1. Introduction

In this article, we present the Linguacuisine app, which was a product of the Erasmus Plus-funded project 'Linguacuisine'. The EU project funding ended in October 2018, but the project is still ongoing, using internal funding. The app and a wealth of project materials can be found on <https://linguacuisine.com>. Linguacuisine is the third generation of digital technology we have produced, the first two being the French and European Digital Kitchens. Rather unusually, we start the introduction with a comic, which provides an introduction to the Linguacuisine concept and procedures in graphical format, and use comics to illustrate concepts throughout.



Figure 1. Linguacuisine app.

Linguacuisine tackles the universal problem of classroom language teaching, namely that students are rehearsing using the language in classrooms, rather than actually using the language to carry out real-world actions. It also tackles the difficulty of bringing the foreign culture to life, and the issue of how to motivate people to learn languages. A significant challenge for nations worldwide is how to improve the foreign language proficiency of its workforce and students. In countries like the UK, the number of students gaining a qualification in a foreign language has decreased significantly, so the question is: how can we engage people with language learning?

At Newcastle University, a group of linguists and computing scientists have been working together for the last 10 years on what language learning might look like if we asked what young people today are interested in as our starting point. Clearly, they are interested in using digital technology, in overseas travel, in global cuisine and cooking, in hands-on experiences and doing things. We used these interests as the design basis for our solution. Many technological approaches to language learning involve learning in a virtual, online world, but we wanted to use language to

carry out a real-world, practical, engaging task with a tangible end product. We chose cooking as it's a universal physical activity which has considerable resonance with both language and culture. It's so enjoyable that countless TV programmes are devoted to it! It involves all five senses, you can work with friends and eat the end product. But what can you learn while cooking? We found you can learn aspects of a foreign language and culture, as well as digital skills while cooking. But why would anyone want to learn a foreign language while cooking? Because of the intimate connections between language, cuisine and culture. If you think of your favourite festival in your own country, then there will be particular food and language associated with it, which will give a direct window into the culture. Ayeomoni (2011, p. 51) suggests that "the relationship among language, food and culture in a society is an inextricable one". Many adult learners are motivated to learn languages through their interest in foreign cuisine and culture, and this project taps into this motivation. Also, many people find technology an inherently motivating tool for learning, as evidenced by the vast range of digital materials available for learning via a variety of platforms. We also found that you learn foreign words better when you are physically touching food and cooking utensils and using them to prepare food. When you are cooking, you involve all of your senses in the learning experience – touch, smell and taste as well as hearing and seeing.

2. Pedagogical principles

This section (based on Seedhouse, 2017) explains the pedagogical principles underlying Linguacuisine systems, materials and procedures. The pedagogical design is based on the principles of Task-Based Language Learning and Teaching (TBLT) (Ellis, 2003). Tasks are divided into 3 phases: *pre-task*, *during-task* and *post-task*, providing a clear design structure for materials, for conduct of sessions and for evaluation of performance. Seedhouse (2017) demonstrates how the phases are implemented in practice by Conversation Analysis of interactional transcripts of learners working through the cycle. It is argued that the project realises some of the advantages of TBLT using digital technology in a real-world setting outside the classroom. This section explains how the concepts of TBLT were operationalised in the Linguacuisine app.

2.1. What is Task-Based Language Learning and Teaching (TBLT)?

The pedagogical design of Linguacuisine employs TBLT, a well-established approach to language learning which prompts learners to achieve a goal or complete a task (Skehan, 1998, 2003). TBLT seeks to develop students' language through providing a task (such as asking for directions) and then using language to solve it. According to Ellis (2003, p. 9) the criterial features of a task are that: a task is a workplan; meaning is primary (language use rather than form); a classroom task relates directly to real world activities; a task can involve any of the four language skills (speaking, listening, reading and writing); tasks engage cognitive processes; task completion is a priority and assessment is done in terms of outcomes. Samuda and Bygate (2008, p. 7) see TBLT as involving holistic activity in that all sub-areas of language are employed to make meaning. They argue that it is in such holistic language work that key language learning processes take place. It is generally assumed (Ellis, 2003, p. 263) that tasks are carried out in pairs or small groups in order to maximise interaction and autonomy.

There has been a substantial programme of research in relation to TBLT, summarised in Skehan (2003). From the perspective of the Linguacuisine project, the major advantages of TBLT as pedagogy were the following. There was a natural match with the chosen activity of cooking, which could be easily conceptualised as a task, as described above. TBLT has well-developed procedures and principles for task design which could be followed and which blended well with HCI (Human-Computer Interaction) design principles. Johnson (2003, p. 96) stresses the importance of an iterative development cycle when designing language learning tasks. He examines the cyclic episodes that task designers actually go through, listing actions such as 'compare'; 'evaluate'; 'reject'; 'modify' and 'review'. This iterative cycle is very much in harmony with the user-centred design cycle used in pervasive computing and HCI. Dix et al. (2003) specify the cycle as 'identify needs, analysis, design, prototype, evaluate, implement, deploy and recycle'. It therefore proved easy to integrate pedagogical and technological design from this perspective. Tasks form a useful basis for designing research as well as pedagogy.

TBLT has so far predominantly been based on tasks to be undertaken within the classroom which simulate real-world tasks. Some innovations in TBLT have combined language learning with other, non-linguistic skills in a similar way to this project. Paterson and Willis's (2008) *English through Music*, for example, aims to help children to absorb English naturally as they enjoy making music together. However, there have been few attempts to employ TBLT in naturalistic settings outside the classroom; the project described here is innovative in combining TBLT and digital technology in a naturalistic kitchen setting outside the classroom. Whereas classroom-based TBLT may engage the learners' senses in terms of sight, sound and touch, Linguacuisine also engages the senses of smell and taste as well, delivering a vivid, kinesic language learning experience.



Figure 2. Linguacuisine – learning with all your senses.

In relation to TBLT and digital technology, Thomas and Reinders (2010, p. 7) refer to the relative dearth and 'marginalization' of CALL research on tasks. Their collection tackles this issue by identifying and developing a range of areas involving technology-mediated tasks; these are reviewed in chapter 2 of Seedhouse (2017). The Linguacuisine project therefore contributes to the research agendas of both TBLT and technology-mediated TBLT.

2.2. The principles of TBLT and Linguacuisine design

The overarching main cooking task in the kitchen was designed according to Ellis's (2003) criterial features quoted in section 2.1 above, in the following ways: we designed it to encourage learners to focus on meaning rather than purely language – that is, they use the language to complete a culinary task, rather than focusing primarily on the language itself. Secondly, learners must employ all four language skills in a holistic manner to achieve the task. Thirdly, the task is situated in an authentic real-world context, namely the kitchen. The task is goal-oriented, involving the production of a dish. Fourthly, cooking tasks are carried out in pairs. In some cases, this generated interaction in L2. In the UK context, for example, we paired foreign learners of English who did not share an L1, compelling them to communicate in English L2. Finally, learners can measure their own success by non-linguistic goal completion, through cooking and consumption of the food. A further characteristic of the Linguacuisine task is that it is a focused task, in that it is necessary for learners to recognise the spoken form of named L2 vocabulary items in order to carry out the task. Learners are pushed to use these items in L2 talk with each other, but are not compelled; Ellis (2003, p.17) notes that learners can always use communication strategies to avoid using the target feature. Ellis (2003, p.142) suggests that focused tasks are of value

because they involve both reception and production and provide a means of teaching language items communicatively, under real operating conditions.



Figure 3. Linguacuisine – explanations of how the system works.

Ellis (2003, p.21) provides a systematic framework for describing the design features of tasks, in which one must specify the goal, input, conditions, procedures and predicted outcomes. These are applied to the Linguacuisine task as follows:

Table 1. Task Design Features of Linguacuisine	
Goal	<ul style="list-style-type: none"> To cook a meal following L2 instructions; To learn a vocabulary set related to tools, materials and processes; utensils, ingredients and cooking processes.
Input	<ul style="list-style-type: none"> L2 spoken, written, video and graphical input provided by the Linguacuisine system; Contextual information is provided by the kitchen environment.
Conditions	<ul style="list-style-type: none"> This is a convergent task in that users must agree on how to cook the meal and a single outcome is targeted. All users receive the same basic information, but receive individualised feedback according to their choices and task progress.
Procedures	<ul style="list-style-type: none"> The task is intended for pairwork and for users to collaborate and produce some L2 talk related to cooking procedures.
Predicted Outcome	<ul style="list-style-type: none"> A meal from the L2 cuisine which can be eaten. Linguistically, it is predicted that some specific L2 vocabulary items will be learnt. Specifically, there will be concrete items (e.g., utensils and ingredients) manipulated during the task.

2.3. Phase framework

In order to operationalize TBLT in this setting, we adopted the cyclical pedagogic TBLT framework put forward by Skehan (1998) and Ellis (2003), which divides activity related to the completion of a task into 3 phases: *pre-task*, *during-task* and *post-task*. This provided both a clear design structure for materials and a guide to implementation. The *pre-task* functions as a preparation stage for the main activity to be carried out in the *during-task* phase. The *during-task* phase involves the performance of the main task set. The *post-task* phase is designed to manipulate attention through reflection on and analysis of *during-task* performance, identification of what has been learnt, and as a period of evaluation of the task outcomes.

2.3.1 Pre-task

The *pre-task* functions as a preparation stage for the activity to be carried out in the *during-task* phase. This may include the presentation of new language, the mobilisation of existing language knowledge and clarification of the type of knowledge that would be required (Skehan, 1998, p. 138). All three features directly relate to preparing or priming the learners' attentional resources and are based on the operations involved in processing information in the short-term and working memory. The learners should get an indication of the purpose of the task and the kind of task it is. The *pre-task* in *Linguacuisine* involves a dual focus on cooking and L2 skills and is divided into presentation and preparation of the L2 items and cooking. Firstly, learners could (where available) watch a purpose-made video recording with optional sub-titles of a native-English speaker making the chosen dish for the project, *English Scones*. This familiarised them with both the cooking procedures required and with the English language to be employed. This facility enabled individualisation of learning. In TBLT terms, this aspect of the *pre-task* framed the main task, motivated the learners and focused their attention on the L2 words which they would encounter during the main task. It introduced them to the process by which they would generate the task output, namely the dish.

Secondly, the learners were able to see photos of the different utensils and ingredients they would need to make the dish and hear their names pronounced in the L2 via an audio file, in order to familiarise them with the specific L2 vocabulary required for the task this is 'list all ingredients/utensils' on the interface. This introduced new language and mobilised existing resources.

Thirdly, seeing photos and listening to audio files of the different utensils and ingredients also constituted instructions to locate these items and prepare them for cooking. So, whereas the first two *pre-task* elements were passive and involved listening, the third element involved learners in actively preparing equipment and ingredients. This element focuses them both on the language required and on the physical materials required for the cooking.

The role of the *pre-task* in the overall cooking session is to prepare the users for the cooking activity. Its pedagogical aim is to provide input about cooking and language through the notions of *preparation* and *presentation*. In TBLT, these introduce learners to the linguistic and procedural knowledge required to complete the task. In *Linguacuisine* we re-specified the notions of presentation and preparation to a dual focus on language and cooking. In TBLT terms, the *pre-task* obliges users to notice and process specific vocabulary items in the input. The content of the feature is provided by the requirement to locate and move the object itself onto the work surface and the linguistic form is salient as it is supplied by the system several times in both spoken and written forms.

2.3.2 During-task

The *during-task* phase involves the performance of the main task set. The *during-task* phase of course entails cooking the dish. It involves step-by-step instructions on how to prepare the dish, together with a range of relevant help. The instructions are verbally communicated by the app as and when required by the learners, using the app interface by pressing relevant buttons. The

cooking task instructions are formulated in such a way as to include cooking-specific vocabulary on which we expect learners would focus most of their attention, having been introduced to the items in the pre-task. The learning environment provides a range of possible supports or scaffolds to cater for a variety of learning styles and L2 proficiency levels, and learners can decide for themselves which to make use of. Videos, photos and audio are available, as well as instructions as written text.

2.3.3 Post-task

The *post-task* phase is designed to manipulate attention through the analysis of during-task performance and reflection, as a period of evaluation and consolidation after the completion of the task. It can also involve identification of what has been learnt, and evaluation of the task outcomes. Skehan presents the post-task as an alternative to what he calls “within-task interference”, that is the disruption that might be caused to the preservation of the communicative purpose if learners were too focused on attention to language features in the performance of the during-task phase (1998 : 148). This is similar to the ‘plenary’ section of a school lesson where a teacher goes through the learning objectives of a lesson and pupils identify ‘what they have learned’. The *post-task* in Linguacuisine focuses on evaluation of what the users had learnt, as well as sampling of the task outcome, namely the dish produced. Targeted vocabulary can be revisited by the learners through looking at the equipment and ingredients again on the app and checking their L2 names. So, whilst the focus *during-task* was on meaning and task completion, the focus *post-task* can be partly on linguistic form and on the language used, as well as on the dish itself. Moreover, the post-task phase provides an opportunity for reflection and discussion.

There are also other possible post-task activities which may be added to the app under ‘extras’. Films or pdf files relating to culture, history, language and cuisine may be added for supplementary use in the post-task phase. A good example of a film is about Italian regional cuisines and can be found with the Italian recipe ‘Involtini’. A good example of a supplementary non-digital activity is provided on <https://linguacuisine.com/larkspur-primary-visit-july-2019/> in which primary pupils produced paper menus in French.

2.4. Relating the principles of TBLT to the Linguacuisine tasks

Ellis (2003, p. 276) introduces eight principles of TBLT which can be used to guide implementation and design of participation. In this section, we see how these were implemented in relation to Linguacuisine.

- *Ensure an appropriate level of task difficulty.* This was implemented by having a wide range of available resources (recipes) and an optional introductory video with a range of options, so users could tackle the task by choosing the resources suitable to their own level.

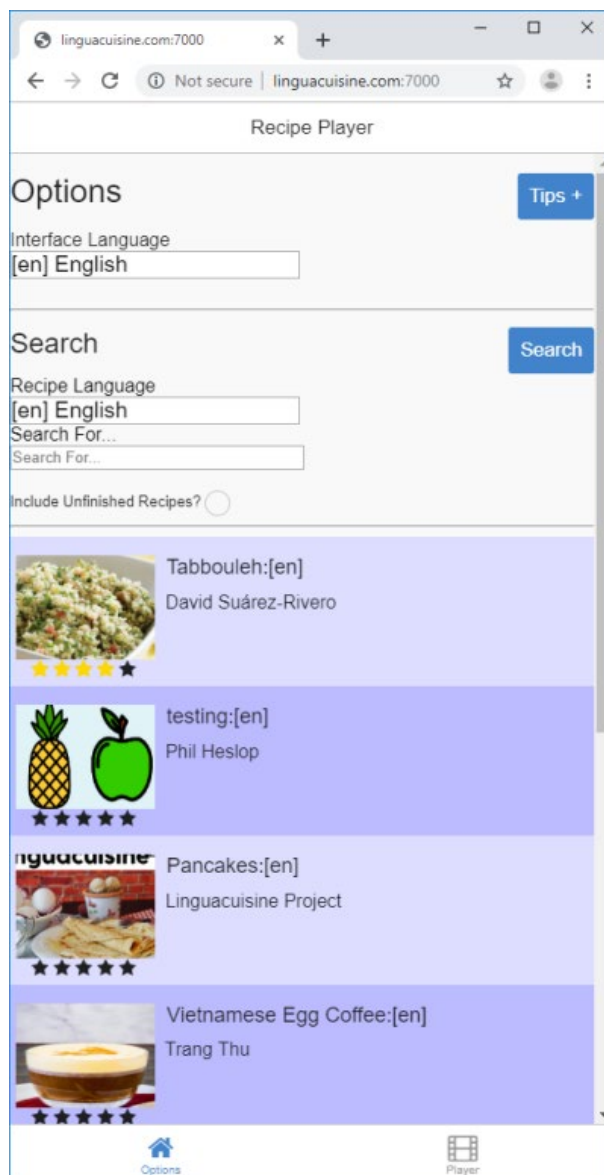


Figure 4. Linguacuisine recipes.

- *Establish clear goals for each task-based lesson.* The main goal of cooking a dish was implemented by showing the video of the dish being prepared, including the final result. Goals for vocabulary learning were established in the pre-task by introducing the target items in both photo and audio formats.

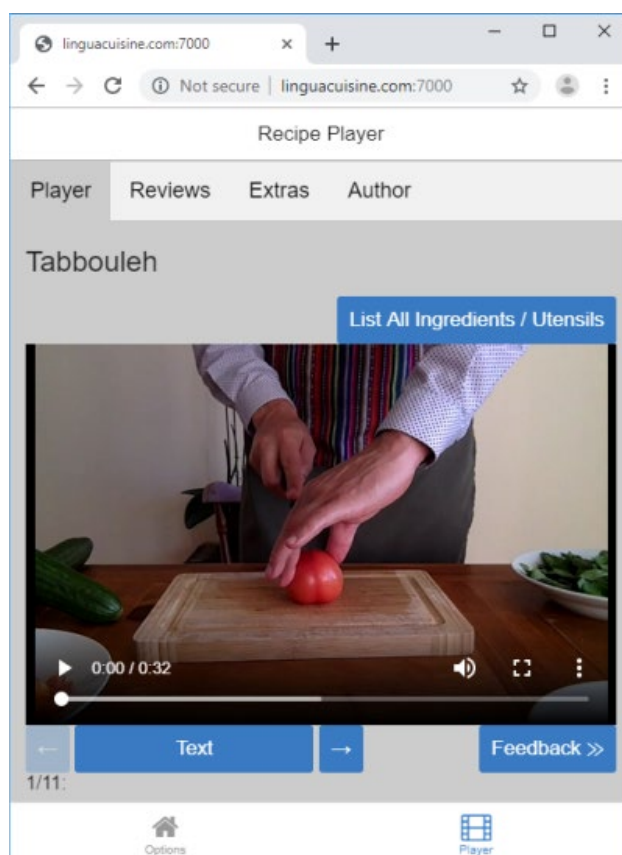


Figure 5. Video of a dish being prepared.

- *Develop an appropriate orientation to performing the task in the students.* This was developed by supplying information about the task in advance of the session to users, by preparing them for the task in the pre-task and reflecting on it in the post-task.
- *Ensure that students adopt an active role in task-based lessons.* The system was designed to require the users to take decisions and perform physical actions on their own initiative. There is normally no teacher present, although there can be if required.
- *Encourage students to take risks.* Users are told that they can make their own decisions as to which resources to make use of in order to complete the task.
- *Ensure that students are primarily focused on meaning when they perform a task.* Users must focus primarily on carrying out the physical task by manipulating utensils and ingredients.
- *Provide opportunities for focusing on form.* Users are able to summon help when they have problems in understanding L2 instructions. The help facility provides help in the linguistic form of the L2 target item in both spoken and written forms.

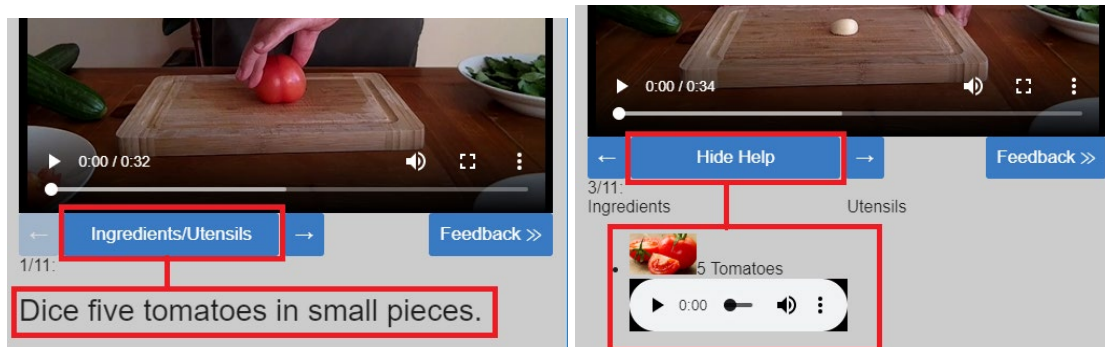


Figure 6. Help provided with linguistic form.

- *Require students to evaluate their performance and progress.* In the post-task users may reflect on and evaluate their task completion and their learning.

So, it has been possible to implement TBLT principles and procedures in the design and implementation of tasks for Linguacuisine.

2.5. Digital competency

As well as a language learning purpose, the apps were designed with participants to improve their digital competency. In order to create a coherent recipe, authors need basic video editing skills, need to know how to upload files to the internet and need to understand some of the underlying technological structure of a recipe (steps, ingredients, utensils and extras). Working with a group of digitally marginalised participants during the design phase, we tested the participants pre- and post-designing and using the app on their disposition towards technology using questions taken from the digital competence framework (Carretero et al., 2017).

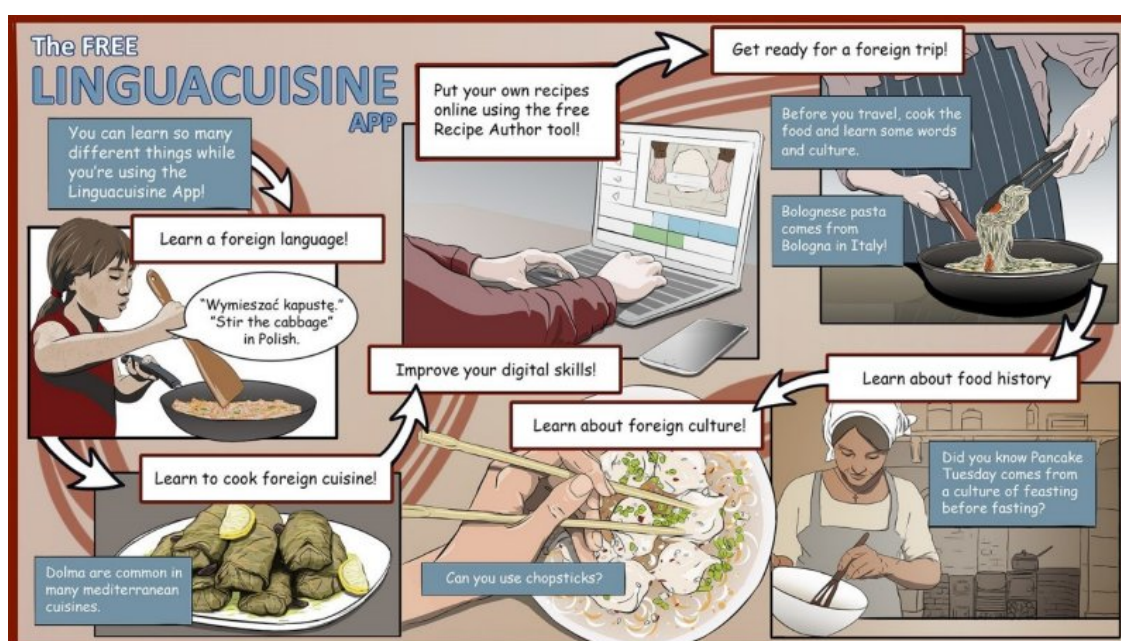


Figure 7. Linguacuisine – supporting different competences.

3. Findings

In this section, we present empirical findings in relation to: firstly, digital competencies and attitudes; secondly, vocabulary learning.

3.1. Findings on digital competencies and attitude (design cohort)

We tested the design cohort (digitally marginalised participants) with a pre- and post-questionnaire about their digital competencies and their general attitude to the learning process:

	Pre	Post	T. Test
Information & digital literacy	2.50	2.84	0.013
Information & data literacy	2.19	2.72	0.004
Communication & collaboration	2.18	2.64	0.017
Digital content creation	1.84	2.51	0.001

Average	2.18	2.68	
Standard Deviation	0.745	0.724	

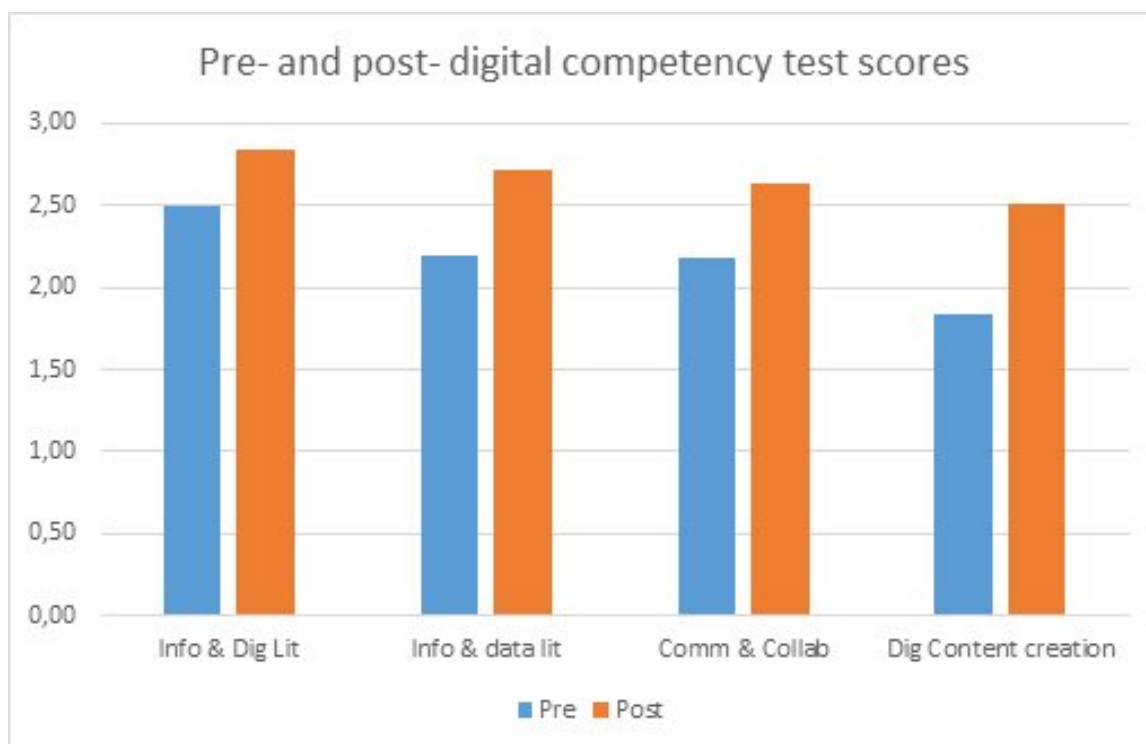


Figure 8. Digital competencies results.

The Digital Competencies questionnaire (Table 2 & Figure 8) showed significant changes in all areas. This reflected the participants' perception of how they felt they had improved on key digital skills. There was no objective assessment of whether skill was measurably improved.

Table 3. Attitude questionnaire results			
	Pre	Post	T.Test
Anxiety regarding the use of Digital Technologies	1.75	1.63	0.215
Attitude to Foreign Language & Culture	1.60	1.66	0.528
Attitude to using Digital Technologies	1.99	1.97	0.882
Motivation for acquiring digital competences	1.66	1.63	0.723
Average	1.75	1.72	
Standard Deviation	0.527	0.412	

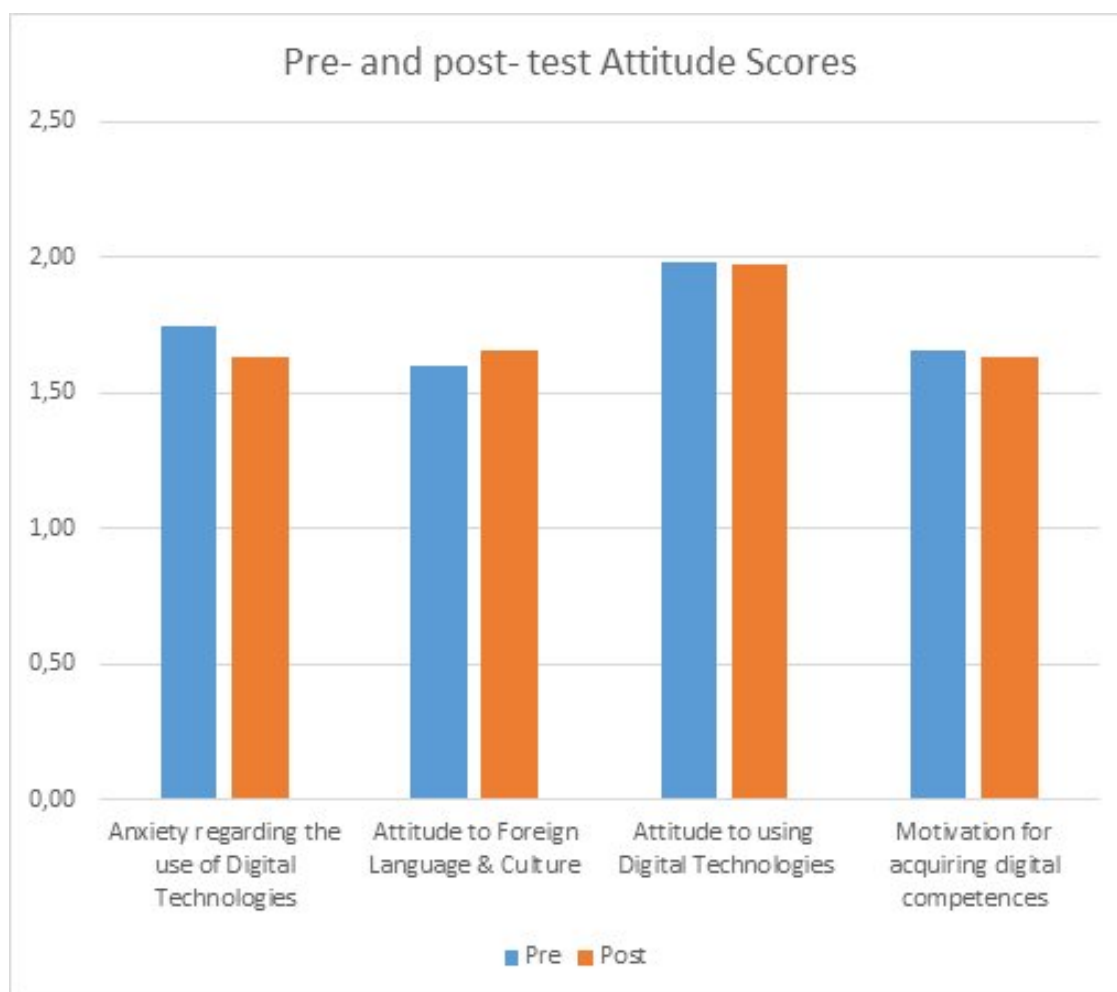


Figure 9: Attitude results.

The Attitude questionnaire (Table 3 & Figure 9) did not show significant improvements in attitudes, although the trend is towards improvement. To understand this aspect of the study, a deeper qualitative thematic analysis was undertaken on participant interviews.

3.2. Findings on vocabulary learning: two studies

One problem for any holistic environment for language learning is how to assess language learning precisely (Seedhouse, 2017). A pervasive digital language learning environment is intended to be a holistic one, in which learners autonomously access resources to complete a task and thereby learn aspects of a language as well as other skills. However, this does pose certain problems when it comes to the precise evaluation of the learning effectiveness of such an environment. Exactly which aspects of a language have been learnt? How do we know participants did not know an item previously and what is the evidence that it has now actually been learnt? More generally, if we are trying to create an autonomous, holistic environment, would this not be disrupted by testing procedures? Ideally, the evaluation of a holistic environment would itself be holistic and evaluate all aspects of language learning together. In Seedhouse (2017), for example, we provided a holistic illustration of learning processes in the French and European Digital Kitchens by presenting representative episodes from a complete task cycle. In this article, by contrast, we decided on a narrow focus on one specific component of language learning for evaluation. This would enable us to see whether there was concrete evidence of learning in one narrowly delineated component of the overall language learning system. The main research question was: to what extent does learners' ability to verbally produce specific vocabulary items change as a result of a cooking session in this pervasive digital environment? The basic research design (described below) was a pre- test/ post- test of specific vocabulary items, carried out on

72 learners of Chinese in China and 24 learners of Vietnamese in the UK. The intervention which was intended to promote learning of the items was the complete experience of a cooking session using the Linguacuisine app, lasting about an hour.

Both the Chinese and Vietnamese studies followed the same basic research design to determine evidence of L2 vocabulary learning.

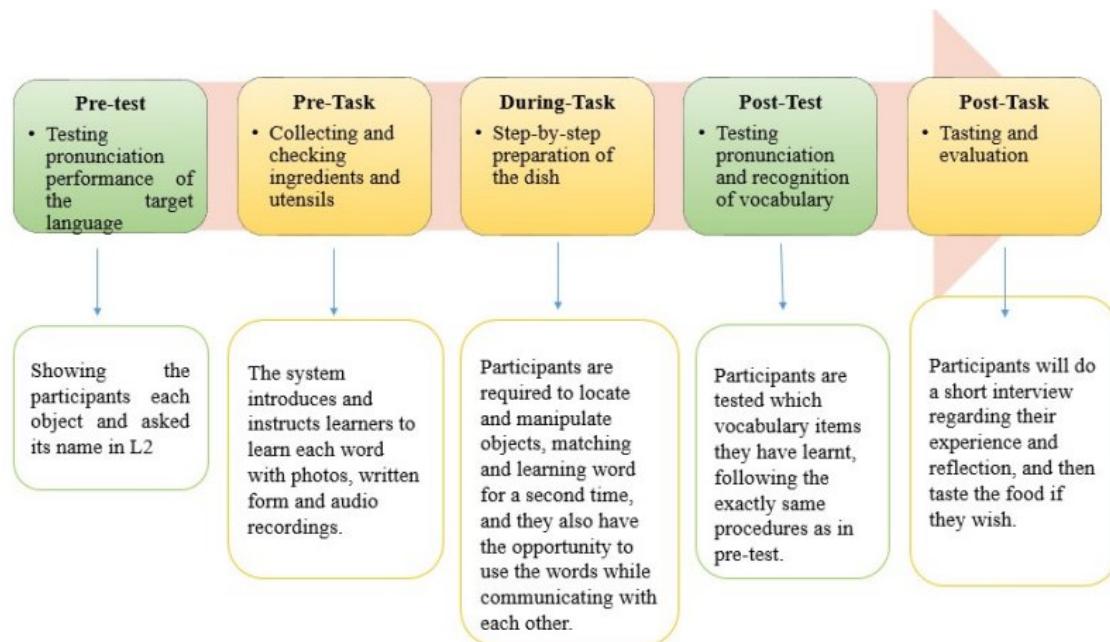


Figure 10. The tasks and tests cycle.

According to Figure 10, we showed the testee each object in order and asked its name in L2, using an audio recorder to record what they answered, if anything, for each item. We therefore established the extent to which each individual was able to actively produce each item prior to their cooking tasks, using the rating scale in Table 4. After they finished their cooking task, each individual completed the post-test immediately and also separately following the same procedure as the pre-test. This could therefore enable us to record granular evidence in terms of individual changes in active production of learning the specific vocabulary items during their completion of tasks.

The rating scale employed in the present study was an adaptation of the Lexical Production Scoring Protocol-Written (LPSP-Written) (Barcroft, 2002) as shown in Table 4.

Table 4. Rating scale for Linguacuisine vocabulary test	
Score	Speaker Spoken Production
0.00 points	The speaker says nothing at all or states that s/he is unable to answer.
0.25 points	The speaker makes an attempt to name the target object which is unintelligible and is very difficult to understand in relation to the target object.
0.50 points	The speaker produces the target lexical item partially , or in a way which can only be understood to relate to the target object with some difficulty, with a major problem in pronunciation and/or clarity. Or the speaker tried to describe the object rather than name it.
0.75 points	The speaker produces the entire target lexical item in an intelligible way, but with a minor problem in pronunciation and/or clarity, or in delivery.

1.00 points	The speaker produces the entire target lexical item with precision and clarity.
-------------	---

3.2.1. Chinese vocabulary learning study

There were 72 international students of L2 Chinese resident in Xi'an, China, where the present study was conducted during March to May 2019. All the participants were assessed on the same 27 vocabulary items (related utensils and ingredients) on 2 occasions (pre-test and post-test) with 5 rating options. The recipe was a traditional Chinese recipe: Eggplant Stir Fry. Participants were 43 males and 29 females in total, age ranged from 18-40 years old, and their exposure to Chinese varied between 2 months and 68 months (5 years and 8 months), with a mean of 13 months (1 year and a month). We tried to pair the participants so that one had a higher language proficiency than the other.

In most cases, the two participants did not have a common L1 and spoke Chinese L2 or English L2 (beginners level participants were allowed to speak English due to their limited language abilities in Chinese) the whole time. Participants who had a common L1 were requested to speak Chinese L2 / English L2, whereas it happened sometimes that they spoke a mixture of Chinese and English.

Table 5. Background information of participants

Kyrgyzstan (n=6)	Surinam (n=1)	Italy (n=1)	Uzbekistan (n=5)
South Korea (n=4)	Kazakhstan (n=12)	Sudan (n=1)	Pakistan (n=19)
Ukraine (n=1)	Japan (n=3)	Morocco (n=1)	Norway (n=2)
Russia (n=3)	Tajikistan (n=5)	Benin (n=1)	Belgium (n=2)
Nigeria (n=1)	Mauritania (n=1)	Turkmenistan (n=1)	Afghanistan (n=2)

In order to find the participants' language proficiency level, their HSK (HanYu ShuiPing KaoShi) test results were established. This is the standardized test of Standard Chinese Language Proficiency of China for non-native speakers such as foreign students and overseas Chinese (see Appendix for HSK to CEFR Description). There were 18 participants at beginner (HSK 1-2), 26 at intermediate (HSK 3-4) and 28 at advanced level (HSK 5-6).

In all cases, to determine the changes in participants' learning outcomes between the pre-test and post-test, we ran a t-test. In the present study, the null hypothesis is that the mean score of pre-test minus the mean score of post-test is equal to 0, which means there is no significant difference between pre-test and post-test. The alternative hypothesis is that the difference in means is not equal to 0, which indicates there is a significant difference.

As shown in Table 7, we found that the t-statistic is 5.581 and p value is < 0.05. The larger the absolute value of the t-value, the smaller the p-value, and the greater the evidence against the null hypothesis. The null hypothesis was therefore rejected and we accepted the alternative hypothesis. Furthermore, the mean score of the pre-test is significantly smaller than the mean score of the post-test.

Figure 11 shows the pre-test and post-test scores for an aggregation of the whole cohort and all the individual lexical items. The horizontal number, from 1-27, represents the exact same order as in the test. The vertical axis number gives the mean test scores for each individual item using the rating scale showed in Table 4, therefore, the minimum score is 0 when participants said nothing at all or stated he/she was unable to answer. By contrast, the maximum goes to 1 which represents that the participants produced the entire lexical item with precision and clarity.

Figure 12 and Table 6 show that the mean score of an aggregate of all the 27 items for the entire cohort rose from 10.465 in the pre-test to 16.872 in the post-test. These differences were all statistically significant (See Table 7). The vertical axis in Figure 8 shows the maximum score would be 27.

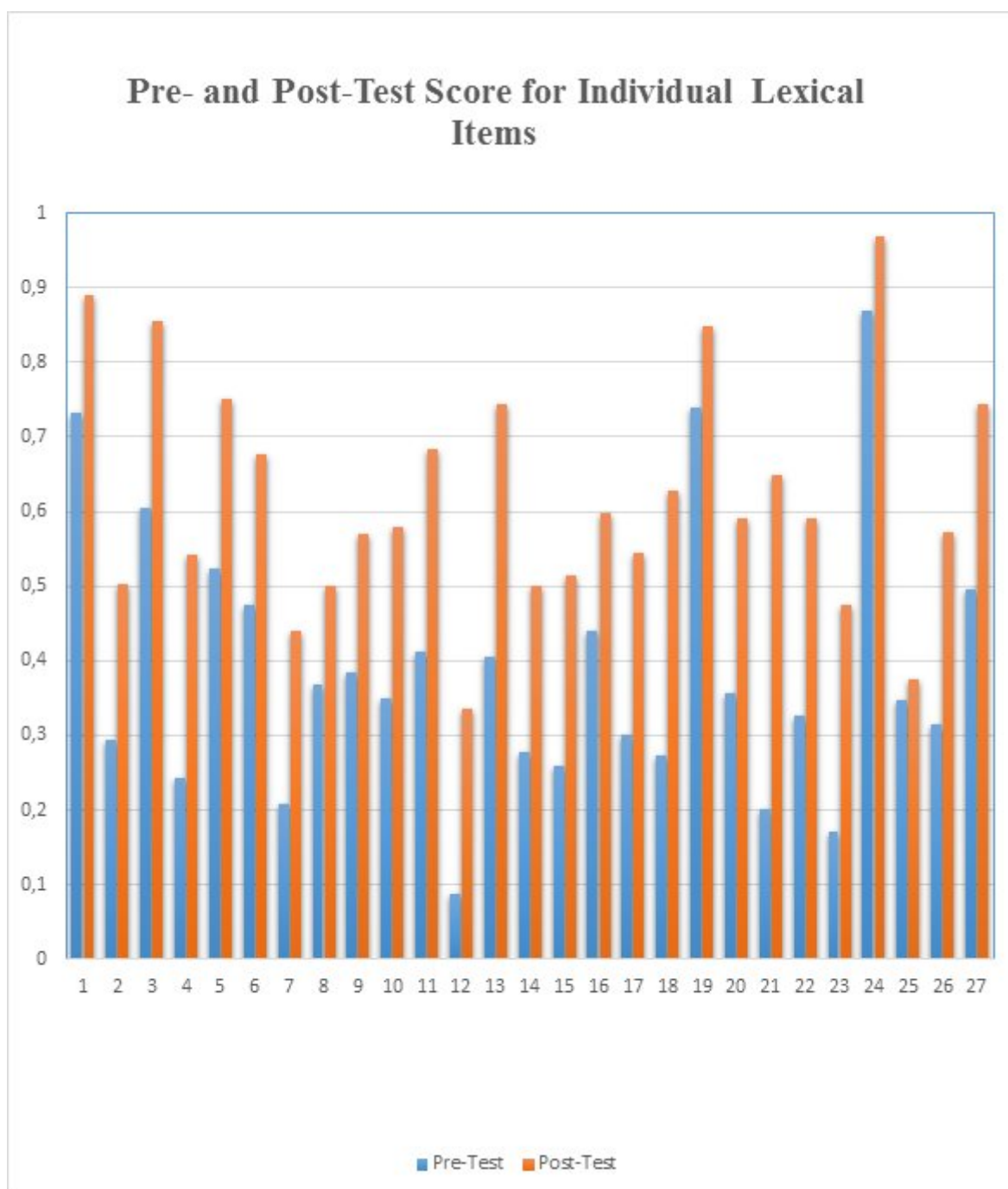


Figure 11. Pre-test and post-test scores for individual lexical items in *Chinese Digital Kitchen*.

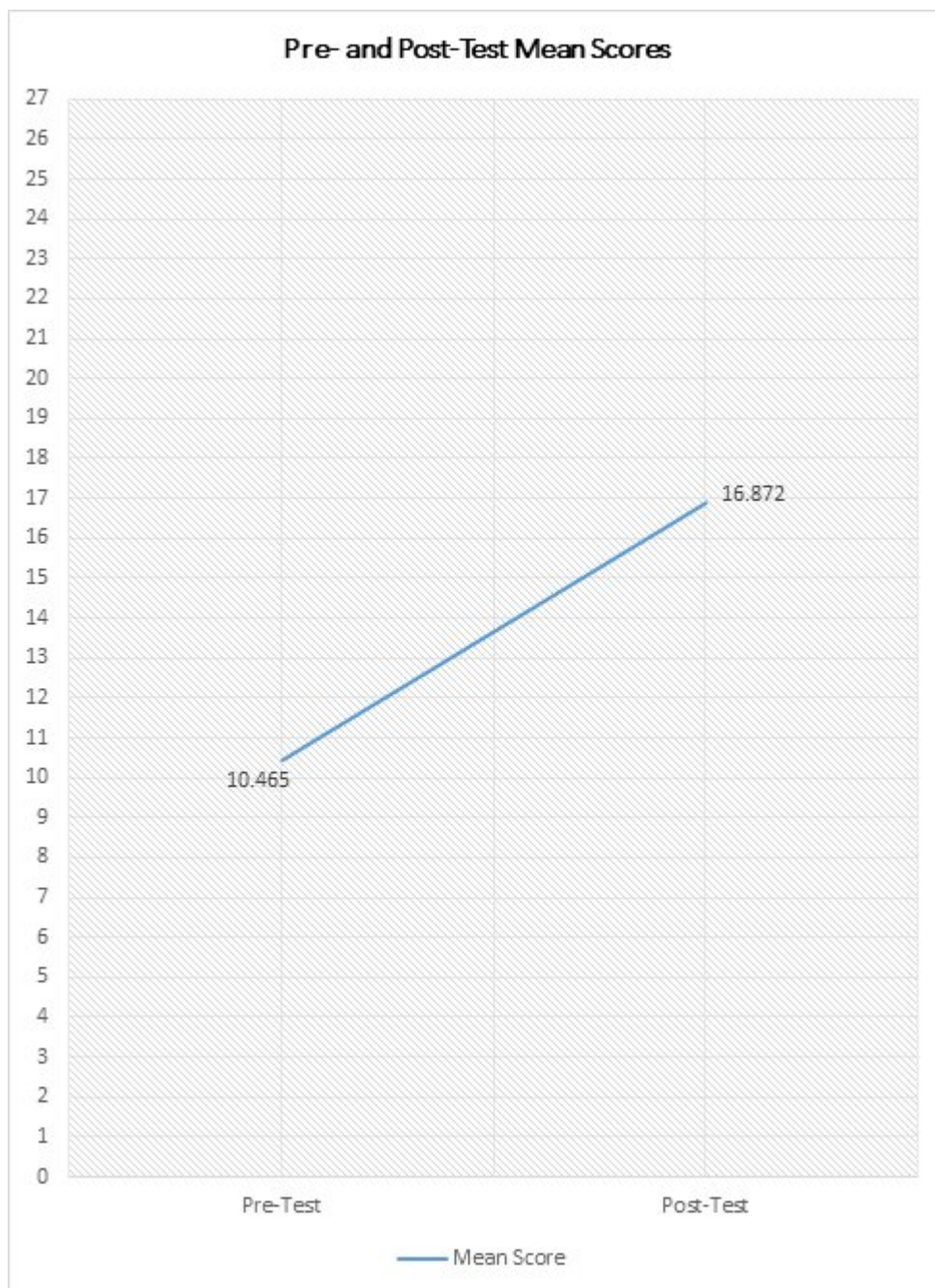


Figure 12. Mean scores on 27 items for the whole cohort in pre- and post-test in Chinese Digital Kitchen.

Table 6. Mean scores and standard deviation			
Overall Score	Mean	Std. dev.	N
Pre-Test	10.465	6.736	72
Post-Test	16.872	7.031	72
Improvement	6.407	3.465	72

Table 7. Statistical significance in relation to the tests			
		P Value	T Value
Pre-Test	Post-Test	< 0.05	5.581

Therefore, we can conclude that in the current study, while working with the international university students learning Chinese, there was a significant gain in the mean score between pre-test and post-test for these items when aggregated. The task-based protocol employed in the present study worked effectively, in that the pre-task phase made participants become aware of their lexical gap and their need to focus on form.

The degree of gain for individual items showed considerable variation; and there is a *prima facie* case that this variation was related to the degree of prior knowledge of the vocabulary item, although other influences cannot be excluded.

3.2.2. Vietnamese vocabulary learning study

Twenty-four participants who had no prior knowledge of Vietnamese language and culture were selected for this study. They were Newcastle University students (7 undergraduates and 17 postgraduates) with 7 different nationalities (Table 8). There were 4 male and 20 female participants, whose ages ranged from 20 to 32, as shown in Table 8. These individuals were randomised into pairs by using “permuted-block randomization” into four groups of six.

Table 8. Sex and age of participants

Sex	Number	Percentage	Minimum age	20
Females	20	83.3	Maximum age	32
Males	4	16.7	Mean	23.9
Total	24	100.0	Standard deviation	2.76

Table 9. Nationality of participants

China (<i>n</i> =13)	Romania (<i>n</i> =1)
Indonesia (<i>n</i> =3)	India (<i>n</i> =1)
Singapore (<i>n</i> =3)	British (<i>n</i> =1)
Malaysia (<i>n</i> =2)	

The recipe used in this study was “Vietnamese Egg Coffee” (Figure 13). Five vocabulary items of utensils and 5 vocabulary items of ingredients were assessed. Each individual lexical item was marked based on Table 4.



Figure 13. Vietnamese egg coffee.

In this study, the test results from pre-tests and post-tests underwent one-way ANOVA (analysis of covariance) conducted in Excel program to test for statistical significance and differences between the four independent groups. These quantitative data are helpful to compare participants' vocabulary knowledge prior and post treatment.

Pre-test vs post-test in the *Vietnamese Digital Kitchen*

Twenty-four participants split equally into four groups did pre-tests and post-tests which involved 10 lexical items. These results were averaged out and plotted to Figure 14 and Table 10, illustrating that the post test results significantly increased. Prior to the learning experience, participants were only able to score 0.061. After the learning session, the post-test score was 0.780 for production. The post-test score shows that the participants improved their vocabulary knowledge.

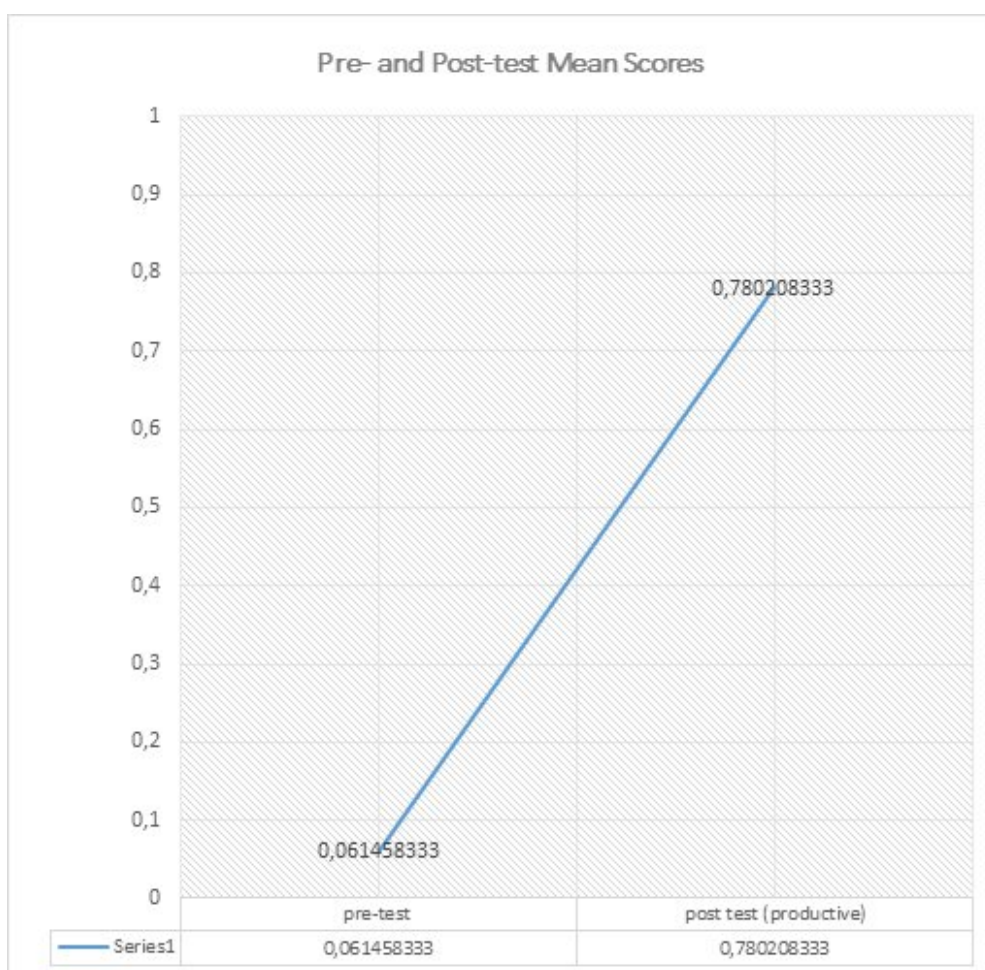


Figure 14. Mean Scores on 10 items for the whole cohort in pre- and post-test in *Vietnamese Digital Kitchen*.

Table 10. Mean score and standard deviation			
Overall Score	Mean	Std. dev.	N
Pre-Test	0.061	0.127	24
Post-Test	0.780	0.013	24
Improvement	0.719	0.121	24

Table 11. Statistical significance in relation to the T-tests			
		P Value	T Value
Pre-test	Post-test	< 0.05	0.001

In Table 11, T-test was used to assess the significance of the post-test results compared to the pre-test results. Productive and receptive post-test results were considered as significant as $p\text{-value} < 0.05$.

3.3. Vocabulary input in the Linguacuisine task cycle

We now consider how the task cycle of Linguacuisine is intended to provide input to vocabulary learning for the learners. The task cycle is separate from the test cycle, although one is wrapped around the other (figure 6). As we saw above, the task structure consisted of pre-task, main task and post-task. In the pre-task, the system introduces the learners to vocabulary items needed in the main task by instructing them verbally to collect the corresponding object from a different area of the kitchen. If the learners do not understand the word spoken by the system, they may call for help in terms of a verbal repetition and a photograph. This ensures receptive recognition of each vocabulary item. Learners therefore have the opportunity to use both the 'guessing from context' and the 'explicit teaching' methods of vocabulary learning (Schmitt and McCarthy, 1997, p. 3). Following its introduction in the pre-task, each vocabulary item is then repeated verbally by the system at least once during the main task (the cooking session) as part of the cooking instructions, thus providing further input. At each point of the cooking session, learners may also request help, which comes in three steps: a repetition of the initial prompt, a picture, while the third consists of a video clip showing the action to be performed. The participants may also produce the vocabulary items when speaking to each another as they conduct the task. The system therefore provides a basis for the learners to both recognize and produce the linguistic form which relates to a specific object. The system requires the learners to physically manipulate the objects during the tasks, while the task design provides the opportunity (but not the necessity) for participants to employ the vocabulary in their joint dialogue.

In the post-task, the participants sample and evaluate the food that they have cooked. This gives them a further opportunity (but not obligation) to employ vocabulary learnt. So, each learner hears the name of each vocabulary item a minimum of two times from the system, but there is no maximum limit. Learners can continue asking the system to repeat the name of an object as many times as they choose, and this particular word may occur an indefinite number of times in their oral interactions.

4. Discussion and conclusions

This article has introduced the TBLT principles which underlie the pedagogical design of the Linguacuisine app, shown how these were operationalised, and illustrated the interactional and learning processes in which learners are engaged. We can conclude that it is indeed possible to employ TBLT principles outside the classroom, and that these provide a suitable basis for designing a digital environment for language learning using an app. We have also shown that testing cycles can be interwoven with task cycles. The two empirical studies of language learning using the app (Chinese and Vietnamese) demonstrate that vocabulary learning gains are significant. The empirical studies of digital competencies showed significant gains in all areas, whereas attitudes showed gains, but not to a significant degree. A wealth of materials and resources related to the app and its use can be found on <https://linguacuisine.com>.

Acknowledgements

The Linguacuisine project was financed by a grant of €324K from the European Union Erasmus Plus Programme 2016-18. Partners: Newcastle University, Hellenic Open University, Workers Educational Association, Action Foundation, University of Modena and Reggio-Emilia.



Figure 15. Linguacuisine recipe authoring software.

References

- Ayeomoni, M.O. (2011). Language, food and culture: Implications for language development and expansion in Nigeria. *International Journal of Educational Research and Technology*, 2(2), 50-55.
- Barcroft, J. (2002). Semantic and Structural Elaboration in L2 Lexical Acquisition, *Language Learning*, 52(2), pp. 323-363.
- Carretero, S. Vuorikari, R. & Punie, Y. (2017). *DigComp 2.1: The Digital Competence Framework for Citizens with eight proficiency levels and examples of use*. EUR 28558. <https://doi.org/10.2760/388>
- Dix, A., Finlay, J., Abowd, G. & Beale, R. (2003). *User-Centred Design*. London: Prentice Hall.
- Ellis, R. (2003). *Task-based Language Learning and Teaching*. Oxford: Oxford University Press.
- Johnson, K. (2003). *Designing language teaching tasks*. Basingstoke: Palgrave Macmillan.
- Paterson, A. & Willis, J. (2008), *English through music*. Oxford: Oxford University Press.
- Samuda, V. & Bygate, M. (2008). *Tasks in Second Language Learning*. Basingstoke: Palgrave Macmillan.
- Schmitt, N. & McCarthy, M. (Eds.) (1997), *Vocabulary: Description, Acquisition, and Pedagogy*. Cambridge: Cambridge University Press.
- Seedhouse, P. (Ed.) (2017). *Task-Based Language Learning in a Real-World Digital Environment: The European Digital Kitchen*. London: Bloomsbury.
- Seedhouse, P., Preston, A., Olivier, P., Jackson, D., Heslop, P., Plötz, T., Balaam, M. & Ali, S. (2013). The French Digital Kitchen: Implementing Task-Based Language Teaching beyond the Classroom. *International Journal of Computer Assisted Language Learning and Teaching*, 3(1), 50-72.

Skehan, P. (1998). *A cognitive approach to language learning*. Oxford: Oxford University Press.

Skehan, P. (2003). Task-based instruction. *Language Teaching*, 36, 1-14.

Thomas, M. & Reinders, H. (Eds.) (2010), *Task-Based Language Learning and Teaching with Technology*. London, New York: Continuum International Publishing Group.

Willis, J. (1996). *A framework for task-based learning*. Harlow, U.K.: Longman.

Editorial Board

Dr. Christine Appel (Universitat Oberta de Catalunya, ES) | Dr. Elisabet Arnó (Universitat Politècnica de Catalunya, ES) | Dr. Elena Bárcena (Universidad Nacional de Educación a Distancia, ES) | Dr. Robert Blake (University of California, Davis, USA) | Dr. Françoise Blin (Dublin City University, IE) | Dr. Alex Boulton (Université de Lorraine, FR) | Dr. María Dolores Castrillo (Universidad Nacional de Educación a Distancia, ES) | Dr. Angela Chambers (University of Limerick, IE) | Dr. Thierry Chanier (Université Clermont 2, FR) | Dr. Letizia Cinganotto (National Institute for Documentation, Innovation, Educational Research, IT) | Dr. Jozef Colpaert (University of Antwerp, BE) | Mr. Anthony Fitzpatrick (Consultant, UK) | Dr. John Gillespie (University of Ulster at Coleraine, UK) | Dr. Marta González-Lloret (University of Hawai'i, Manoa, USA) | Dr. Mar Gutiérrez-Colón Plana (Universitat Rovira i Virgili, ES) | Dr. Trude Heift (Simon Fraser University, Canada) | Dr. Phil Hubbard (Stanford University, USA) | Dr. Kristi Jauregui (University of Utrecht, NL) | Dr. María Jordano (Universidad Nacional de Educación a Distancia, ES) | Dr. Mike Levy (University of Queensland, Australia) | Dr. Peter Liddell (University of Victoria, Canada) | Dr. Antonio Martínez-Sáez (Universidad Autónoma de Madrid, ES) | Dr. Robert O'Dowd (Universidad de León, ES) | Dr. Ana Oskoz (University of Maryland, USA) | Dr. Pascual Pérez-Paredes (Universidad de Murcia, ES) | Dr. David Perry (Universidad Politécnica de Valencia, ES) | Dr. Joan Tomàs Pujolà (Universitat de Barcelona, ES) | Dr. Timothy Read (Universidad Nacional de Educación a Distancia, ES) | Dr. Esperanza Roman-Mendoza (George Mason University, USA) | Dr. Germán Ruipérez (Universidad Nacional de Educación a Distancia, ES) | Dr. Bernd Rüschhoff (University of Duisburg-Essen, DE) | Dr. Shannon Sauro (Malmö University, SE) | Dr. Mathias Schulze (University of Waterloo, Canada) | Dr. Rafael Seiz-Ortiz (Universidad Politécnica de Valencia, ES) | Dr. Ana Sevilla-Pavón (Universitat de València, ES) | Dr. Bryan Smith (Arizona State University, USA) | Dr. Peppi Taalas (University of Jyväskylä, FI) | Dr. Steven Thorne (Pennsylvania State University, USA) | Dr. Margarita Vinagre (Universidad Autónoma de Madrid, ES) | Dr. George S. Ypsilandis (Aristotle University, GR) |

The EUROCALL Review is an open access journal published online biannually by the European Association for Computer-Assisted Language Learning (EUROCALL) and is hosted on the Online Journal System at <http://polipapers.upv.es/index.php/eurocall> by Universitat Politècnica de València (Spain). Regular sections include:

- *Reports on EUROCALL Special Interest Groups: up-to-date information on SIG activities*
- *Projects: reports on on-going CALL or CALL-related R&D projects*
- *Recommended websites: reports and reviews of examples of good practice in language learning website development*
- *Research papers on CALL related topics*
- *Research and Development papers on CALL related topics*
- *Reflective Practice papers on CALL related topics*

The EUROCALL Review is indexed in

- ERIC (Education Resources Information Center - an internet-based digital library of education research and information sponsored by the Institute of Education Sciences of the U.S. Department of Education.)
- DOAJ (Directory of Open Access Journals) - an online directory that indexes and provides access to quality open access, peer-reviewed journals.
- MLA Modern Language Association International Bibliography.
- The European Reference Index for the Humanities and the Social Sciences (ERIH PLUS). This database contains data on scientifically approved European journals within the disciplines of humanities and social sciences.



Red Iberoamericana de Innovación y
Conocimiento Científico

The EUROCALL Review is also listed in the [DULCINEA](#) database, providing bibliographic data, access policies and self-archiving-policies, and in Wiley's [Linguistics Abstracts Online](#) database - a continually updated database with abstracts appearing within 6-12 months of publication.

Articles submitted to *The EUROCALL Review* are peer-reviewed by members of the Editorial Board.

The EUROCALL Review, edited by Prof. Ana Gimeno, is licensed under a [Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License](#). Ownership of copyright remains with the Author(s), provided that, when reproducing the Contribution or extracts from it, the Author(s) acknowledge first publication in *The EUROCALL Review* and provide a full reference or web link as appropriate.

