



European Association for Computer-Assisted Language Learning

THE EUROCALL REVIEW

Volume 26, Number 1, March 2018

Editor: Ana Gimeno

Senior Associate Editor: David Perry

Associate editor: Juan Carlos Casañ Núñez

ISSN: 1695-2618

<http://polipapers.upv.es/index.php/eurocall>

Table of Contents

Research papers

Flipped learning in an EFL environment: Does the teacher's experience affect learning outcomes?
Adrian Leis and Kenneth Brown 3

Digital flashcard L2 Vocabulary learning out-performs traditional flashcards at lower proficiency levels: A mixed-methods study of 139 Japanese university students
Robert John Ashcroft, Robert Cvitkovic and Max Praver 14

Web 2.0 tools in the EFL classroom: Comparing the effects of Facebook and blogs on L2 writing and interaction
Gilbert Dizon and Benjamin Thanyawatpokin 29

Reflective practice paper

Digital video creation in the LSP classroom
Ornaith Rodgers and Labhaoise Ni Dhonnchadha 43

Literature review

Speaking Practice Outside the Classroom: A Literature Review of Asynchronous Multimedia-based Oral Communication in Language Learning
Eric H. Young and Rick E. West 59

Recommended app

Designing and assessing a digital, discipline-specific literacy assessment tool
Paul Graham Kebble 79

Editorial Board 89

Research paper

Flipped learning in an EFL environment: Does the teacher's experience affect learning outcomes?

Adrian Leis* and Kenneth Brown**

*Miyagi University of Education, Japan | **University of Teacher Education
Fukuoka, Japan

* adrian@staff.miyakyo-u.ac.jp | ** kbrownor@fukuoka-edu.ac.jp

Abstract

In this paper, the authors discuss the findings of a quasi-experimental study of the flipped learning approach in an EFL environment. The authors investigated the composition-writing proficiency of two groups of Japanese university students ($n = 38$). The teacher of one of these groups had had much experience teaching with the flipped learning model, whereas the other teacher had had no experience. The first aim of the study was to discover if improvements in writing proficiency could be observed within each group. The results indicated that statistically significant improvements were seen both for students studying under a teacher with experience conducting flipped classrooms, $t(16) = 4.80, p < .001, d = 1.27$, and a teacher without flipped classroom experience, $t(20) = 7.73, p < .001, d = 1.61$. The second aim of the study was to investigate whether any differences in improvement between the two groups occurred. The results suggested that students in both groups improved at similar degrees: $F(1, 36) = .087, p = .77$. These results suggest that regardless of a teacher's experience with the flipped learning approach, it appears to be a successful way of teaching in EFL environments. The authors conclude that, regardless of the teacher's experience with the model, the flipped learning approach is an ideal way to increase the amount of individual coaching possible in the classroom, bringing about more efficient learning.

Keywords: Flipped learning, writing, proficiency, university students, learner agency.

1. Introduction

The increased possibilities for creativity that come with Web 2.0 have generated much interest among EFL researchers and educators, especially in the field of computer-assisted language learning (CALL). The use of blended learning (i.e., integrating the utilization of the Internet into regular classroom environments) has especially grown in popularity over the past decade. The flipped classroom, one example of blended learning, has also received much attention from a wide range of scholars. In the flipped classroom, teachers' explanations of the content of the textbook or grammar points, for example, that would normally be given during class time are provided for students usually through some medium requiring use of the Internet. Because the students have already listened to their teachers explaining what was to be studied during class, more time is available for students to do practice exercises and tasks under the direct supervision of the teacher. With individualized instruction being an essential aspect of education (Keefe, 2007), it is vital that teachers look for ways to increase the possibility of making this a major part of their classrooms. The increased opportunity for personalized coaching which comes with the flipped classroom is perhaps the greatest benefit of this approach to teaching. Table 1 displays a simplified comparison of the learning structures of a traditional class and a flipped class as used in the present study.

Table 1. A diagram comparing the students' activities before, during, and after class in a traditional and flipped classroom

	Before class	During class	After class
Traditional learning	Students prepare for class alone	Students listen to the teacher explain lesson content	Students do practice exercises alone
Flipped learning	Students watch a video explaining lesson content	Students do practice exercises under the teacher's supervision	Students review based on the teacher's advice

In this paper, the authors describe the adoption of the flipped classroom model in two separate classrooms with two different teachers. The first aim of the study was to reinvestigate whether using the flipped learning approach would bring about increases in students' EFL composition writing proficiency. Second, the authors looked at whether there were any differences in the levels of improvement between the two groups. The teacher of one group of students had had a wealth of experience with the flipped learning model and created the videos being used in this study. The other teacher had had no experience with the flipped learning model. Therefore, the authors wished to investigate whether the students in both groups would see similar increases in proficiency. If this could be achieved, it could be concluded that the flipped learning model is indeed effective, regardless of the experience the teacher has had with this approach to foreign language education.

2. Literature review

Although the concept of flipped learning has been around for many years, it has only been in the last few years that it has increased in popularity. In some of the earliest reports on the advantages of providing students with guided class preparation, Mazur (1997) and later Crouch and Mazur (2001) described physics classes in which students were given the teacher's lecture notes one week before the actual lecture. Then, after the students were given a short quiz to confirm their understanding of the content of the lecture notes, most of the class time was spent on discussions, with the lecture notes that had been distributed to students being viewed as guides to assist students as they prepared for each lesson. This approach of providing more guidance as students prepared for class, coined *classroom flip* by Baker (2000) and *inverted classroom* by other researchers (e.g., Lage, Platt, & Treglia, 2000), has allowed for communication between teachers and students to go beyond the constraints of the classroom, thus setting the tone for educators in the early years of the twenty-first century.

In 2012, the idea of flipped learning rapidly increased in popularity as knowledge of the teaching approach spread through various media such as the book *Flip your Classroom* (Bergmann & Sams, 2012) and the online teaching resource, Khan Academy. Since then, there has been a gradual increase in the amount of research focusing on the effects of the flipped classroom in an EFL environment. One of the earlier works by Brinks Lockwood (2014) used videos and lesson materials related to topics studied in class to create a flipped learning environment in her classroom. Using previously prepared materials and publicly available videos reduced the burden on the teacher. In some ways, this answers one weakness of the flipped learning model, that is, the length of time it can take to create materials for the flipped classroom, which has been described in previous studies (e.g., Leis, Cooke, & Tohei, 2015).

Other studies in an EFL environment have looked at changes in classroom atmosphere and students' attitudes as the result of studying under the flipped learning method. Mehring (2015), for example, conducted interviews with students to investigate their opinions of learning English in a flipped classroom. He suggested that more opportunities for active learning were made possible in comparison to those available in traditional classrooms. This was due to the extra time available during the classes that had been used for explanation of the textbook content in traditional environments. Forsythe (2016) had similar positive arguments for the flipped approach, suggesting

that it promoted autonomous learning, with students being free to choose where and when they wanted to listen to and watch the teachers' explanations. Lee and Wallace (2017) suggested students studying English in South Korea under the flipped learning approach liked this way of learning English due to the control they had over how they prepared for class and the increased time to use English during class. Only 10% of students in the flipped classroom of Lee and Wallace's study expressed dislike of this method, claiming it gave the students too much homework and was too difficult compared to the traditional way of teaching. The majority of the students in their investigation said that the flipped classroom was fun and showed gratitude for the efforts exerted by the teachers to make the videos. Leis (2015) suggested similar appreciation had been demonstrated by students in his study. He reported that students were more highly motivated to study harder due to the efforts of their teacher in creating the videos for class: "I understood how hard my teacher worked to create the videos for the class, so I felt a responsibility to work hard with my study." Whether similar efforts would be seen by students studying in a class in which the videos used were not created by their regular teacher will be one of the focuses of the study described in this paper.

Although the amount of literature on flipped learning in the EFL environment is gradually increasing, there is, as suggested by Lee and Wallace (2017), still a lack of studies investigating its effects on students' proficiency. In another study conducted in Taiwan, Hung (2015) concluded that students studying under the flipped learning approach tended to perform better than those who were not. However, the time span for Hung's study was very short (i.e., three lessons in six weeks). This is much shorter than one full university semester (i.e., 15 weeks), which has been suggested as the minimum period for which significant improvements in linguistic performance and motivation can be expected to be seen (Sasaki, 2011).

There are, however, a few studies that have looked at the effects of flipped learning on linguistic proficiency for a full university semester. In Lee and Wallace's (2017) study, statistically significant improvements in proficiency were only observed for the flipped learning group in the end-of-semester tests and one of their writing assignments. One of the reasons suggested for this was that many East Asian students tend to prefer the passive style of learning that is seen in traditional classrooms. Leis (2016), for example, suggested that students' self-perceived proficiency (i.e., linguistic confidence) increased under the flipped learning method. This was because students were able to prepare better for discussions that were going to be held in the class by watching the videos with teachers' explanations. Furthermore, Leis (2015) and Leis, Cooke, and Tohei (2015) provided empirical evidence to support the notion that the use of flipped learning brings about significant improvements in linguistic proficiency in English composition classes in comparison to the traditional classroom. It was argued that combinations of the freedom of when and where to watch the videos, the closed captions used in the videos, and the individual instruction that came with the flipped learning environment influenced increases in proficiency. In the present paper, the authors aim to add further evidence concerning the effects on students' proficiency when studying in a flipped classroom, as well as investigate any differences that may occur depending on the teacher's experience with the flipped learning model.

3. The Study

3.1. Participants

A total of 38 second-year Japanese university students studying at two separate institutions participated in the study. Because they were studying at two different institutions, the authors were unable to create groups of statistically equal proficiency based on a pre-test. Both groups studied English composition under the flipped learning model. Ideally, a comparative investigation of two (or more) English composition classes with one conducted in a flipped environment and the other in a traditional manner would give a more precise indication of whether the flipped learning approach is successful or not. However, in this investigation, neither class was taught in the traditional manner because one of the authors had previously conducted studies comparing the flipped and traditional approaches in English composition classes (e.g., Leis, Cooke, & Tohei, 2015). In all of the studies, the flipped approach proved to be

more successful. Thus, in order to provide what was thought to be a more effective teaching approach, both groups were taught in flipped classrooms. Furthermore, the focus of this study was on whether a teacher who had had little to zero experience with the flipped learning model could see similar improvements in his/her students as a teacher who had had much experience.

Table 2. Descriptives of participants in the present study		NFE	FE
Majors		Education	English
N		21	17
Experience with flipped learning		None	None
Age		19.21	19.25

Table 2 displays descriptions of the participants. The group whose teacher had had experience with flipped learning and created the videos used throughout the course was given the title of Flipped Experience (i.e., FE). The second group, whose teacher had had no experience with flipped learning and had not created the videos to be used was named No Flipped Experience (i.e., NFE). None of the students had had experience of studying in a flipped learning environment. All of the students in the FE group were English majors, and the English composition course used for this study was compulsory for graduation. All of the students in the NFE group took the English Composition course as a requirement for gaining a teaching license.

3.2. Methodology

The present quasi-experimental study was conducted over a 15-week English composition course at two separate Japanese universities. Although, as the authors admit, a quasi-experimental design is not ideal for a scientific investigation, due to the lack of randomness in choosing the subjects for the study, because the distance between the two universities at which participants in the study attended was more than 1000 kilometers, it was not practical to use a pure experimental design. The courses included orientation in the first week, a mid-course review test in the eighth week, and a final review test in the final class. Because the teachers wanted to provide students with an ample input of compositions, a reading textbook was chosen for the course with explanations of the structure of English compositions being given based on passages appearing in the textbook.

In Week 1 of the study, the participants were asked to write compositions for homework based on the theme, "An event I participated in during high school." The students were then given English compositions on various topics throughout the course (e.g., recipes and historical characters). Then in Week 14 of the course (i.e., the week before the final review test), the students were asked to write compositions on the same topic as Week 1. Because the researchers wanted to focus on writing proficiency, rather than writing fluency, compositions were given to students as homework tasks to be completed before class rather than during class under time constraints. Table 2 displays a summary of the recommended weekly study schedule given to students at the beginning of the courses.

Table 3. The weekly preparation schedule recommended to students in this study							
Day	1	2	3	4	5	6	7
Plan	Watch video with captions	Submit vocabulary homework	Watch video with captions	Write first draft	Watch video without captions Check draft	Check draft	Attend class

The teachers of both groups in the study followed a similar lesson plan. First, confirmation and explanations of any difficulties that were observed with homework were given, before a brief summary of the video content that students would have watched before the class. After this, students worked in pairs or small groups to check each other's compositions, creating a peer-coaching environment. During this time, the teacher walked around the classroom, giving individual instruction to students to assist them with improving their composition proficiency. Finally, the teacher gave feedback to students based on any reoccurring difficulties that were noticed during peer-instruction time and made sure students were aware of what they needed to do to prepare for the following lesson. Table 4 shows an example of a summary of the 90-minute lesson structure used by the instructors of the NFE and FE groups.

Time	5	10	10	60	5
Plan	Greetings and confirm class goal	Discuss homework submitted by students	Summary of the video students had watched before class	Students give peer-coaching on essays while the teacher gives individual instruction where needed	Remind students of the topic for the following lesson

Analyses for this study were conducted in a pre-test-post-test design. Three independent teachers of English who were unaware of the purpose of the study gave scores for the Week 1 compositions (i.e., Pre-test) and the Week 14 compositions (i.e., Post-test). The compositions were evaluated using rubrics created by the teacher of the FE group, which were used for feedback throughout the entire course by both teachers. The rubric included five criteria (i.e., introduction, body, conclusion, content, and accuracy), with each having a maximum score of five, giving a maximum total score of 25 for each composition. A Cronbach's Alpha Reliability Analysis was conducted using SPSS Version 23 to compare the scores provided by the three evaluators and was deemed appropriate for statistical analyses (i.e., $\alpha = .95$). Also, in order to gain students' opinions regarding the use of the flipped learning approach, the authors asked students for their views regarding this way of learning how to write English compositions. These responses were not compulsory, and of the 38 participants in the study, the authors received feedback from only five. Details regarding the statistical analysis will be reported in Section 4 of this paper.

3.3. Research questions

In this study, the authors aimed to gain a deeper understanding of the following research questions (RQs):

RQ1. Does the Flipped Learning approach lead to significantly higher proficiency?

RQ2. Are there any significant differences in improvements in proficiency between students studying under the Flipped Learning approach with different teachers who have different amounts of experience with flipped learning?

Based on the results of previous studies (e.g., Leis, Cooke, & Tohei, 2015), the authors hypothesized that the overall English writing proficiency of subjects would increase significantly. Furthermore, it was predicted that similar improvements would be seen by students, regardless of whether the videos being viewed by the students were created by their teacher or not. Although the narration in the videos would have been an unfamiliar voice for the students in the NFE group, as one of the authors suggested before the study, this was not unlike using regular textbooks written by authors unknown to the students. Therefore, even if the teaching materials have not been prepared by the person teaching in the classroom, similar results should be observed if two different teachers are using the same materials in a similar way.

4. Results

The first research question asked whether studying under a flipped learning approach leads to significantly higher proficiency in English composition. Paired-samples *t* tests were conducted to measure the statistical differences between the pre-test and post-

test for the NFE and FE groups. Analyses showed that significantly significant improvements in proficiency with strong effect sizes and no overlaps in the 95% Confidence Intervals (95% CI) at the post-test in comparison to the pre-test had been achieved for both the NFE group, $t(20) = 7.73, p < .001, d = 1.61$, and the FE group, $t(16) = 4.80, p < .001, d = 1.27$. Table 5 and Figure 1 display the composition proficiency scores.

Table 5. Composition proficiency scores in the present study

Group	Test	Score	SD	95% CI
NFE	Pre	9.89	3.79	8.16, 11.61
	Post	16.81*	4.74	15.31, 20.18
FE	Pre	7.31	3.72	5.40, 9.23
	Post	14.29*	6.82	10.79, 17.80

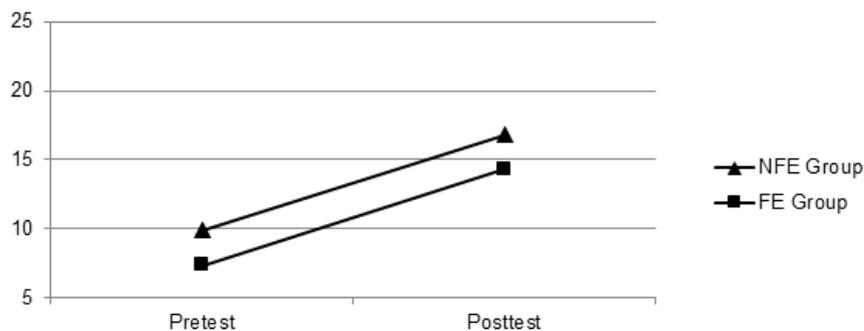


Figure 1. Changes in proficiency of the NFE and FE groups in this study.

In the second RQ, the authors compared the post-test scores for the NFE and FE groups to measure if any differences in proficiency would be seen when studying under different teachers using the same teaching material. First, an independent-samples t test was conducted to investigate if any statistically significant difference existed between the two groups at the pre-test stage. The results were that the NFE group had higher proficiency (i.e., 9.89) at the pre-test stage, $t(36) = 2.10, p = .043$, than the FE group (i.e., 7.31).

Because there was already a statistically significant difference between the two groups at the pre-test, it is possible that any difference observable at the post-test may be due to some factor not related to flipped learning at all. Therefore, conducting an independent-samples t test at the post-test stage may bring unreliable results. To avoid this, the authors used an Analysis of Covariance (i.e., ANCOVA), to compare the data under the presumption of identical scores at the pre-test stage. Larson-Hall suggests an ANCOVA as a statistical analysis to use when “there is some external factor, such as pre-test or TESOL score, which will affect how your students will perform on the response variable” (2010, p. 357). Furthermore, Loewen and Plonsky (2016) suggest that the ANCOVA is often used in SLA research, especially in quasi-experimental designs, as is seen in the present study. This analysis has been used in various studies of second language acquisition (e.g., Fraser, 2007; Larson-Hall, 2008; Lim & Hui Zhong, 2006; Lyster, 2004) in order to remove any differences observed at the pre-test stage. The study by Lim and Hui Zhong (2006), for example, comparing a regular reading class and one conducted in a CALL environment, especially resembles the study described in the present paper as it found statistically significant differences in the pre-test scores. It was therefore deemed satisfactory by the authors of this study to run an ANCOVA to

give an accurate indication of the performance of both groups, even though statistical differences had been observed in the pre-test.

The ANCOVA was conducted with the pre-test scores adjusted to 8.74. The ANCOVA result showed that at the post-test stage there was no statistically significant difference, $F(1, 36) = .087, p = .77$, between the NFE group with an adjusted post-test score of 15.91 and the FE group with an adjusted post-test score of 15.40. Thus, it can be concluded that regardless of the experience the teacher has with flipped learning, there appear to be no major differences in the effects on students' learning.

5. Discussion

The authors of the present paper had two principle objectives in this study: 1) to add depth to the current lack of evidence concerning the effectiveness of flipped learning on improving students' proficiency, and 2) to discover whether students studying under a language teacher who had little to no previous experience with the flipped method would see similar benefits to those studying under a teacher who had had experience and had created the materials that were used in both classes. The results suggest that using the flipped learning method in an EFL composition class assisted in producing statistically significant improvements in students' proficiency, regardless of who was teaching them. The authors will now discuss possible reasons for these results.

First, as mentioned earlier, the availability of explanations of the lesson content online enabled students to view and review the videos at any time and place they wished. This may have promoted autonomy among the students. Autonomy has been defined as "the capacity to take control of one's own learning" (Benson, 2011, p. 58). By giving learners the choice of where, when, and even how they wish to view the teacher's explanation, students in a flipped learning environment may receive such freedom. It has been shown in earlier studies that students' linguistic proficiency may improve when they become autonomous in their learning (Dam & Legenhausen, 1996). However, it is imperative that teachers still be part of this learning process. As Little (1990) explains, autonomy is not merely self-instruction. It still requires input from the teachers. Holec (1981) suggests that for students to be described as autonomous, they need to be taking charge of decisions related to various issues, such as objectives, content, learning techniques, and evaluation. Although flipped learning may not, admittedly, result in students who are autonomous to the degree Holec suggests, it still appears to give students a sense of agency, giving them the feeling that they are in charge of some aspects of their learning.

Second, the use of closed captions and animation, and the capability to re-watch parts of videos that were difficult for students to understand, may have enhanced their comprehension of the content. Japanese students are often described as shy and hesitant to participate actively in language classes, for reasons such as saving face (Harumi, 2011), a lack of self-confidence (Anderson, 1986), and a decline in motivation to study English (Matsukawa & Tachibana, 1996). Therefore, in a traditional classroom environment, it may be unlikely that students would raise their hands to ask teachers to repeat something that had been said or clarify points that had not been understood. The videos used for the flipped classroom described in this study included a variety of features that have previously been shown to improve student listening comprehension, such as closed captions (e.g., Perez, Norgate, & Desmet, 2013; Lee, 2017) and animation and annotation (e.g., Yang & Chang, 2014). Feedback from one student from the NFE Group gave support to the use of closed captions, mentioning that having the closed captions made the videos easy to understand. Because the same videos with the same closed captions were used for both the NFE and FE Groups, students were able to study under the same conditions, thus explaining the similar increases in proficiency of the students.

Third, the flipped learning approach enabled more individual coaching for students by the teachers. In traditional classrooms, teachers waste too much time giving explanations of the textbook (Mazur, 1997). This reduces the number of opportunities for students to receive personalized feedback from their instructors, which would normally allow them to focus on their weaker points and their misunderstandings of the lesson content. In this age, it is no secret that students learn in different ways. Thus

one of the principle goals of teachers should be to give as much attention as possible to each individual student in the classroom (Keefe, 2007). Teachers, however, do need to be aware of the dangers of providing too much care, as students may become reliant on their instructors, forgetting that learning is a lifelong task, and their mentors cannot always be by their sides to help them in tough times. The flipped learning model appears to provide an ideal balance of giving students a sense of agency, thus promoting a degree of autonomy, but at the same time allowing teachers the opportunities to attend to their students' individual needs.

6. Pedagogical implications

Flipped learning appears to have resulted in increased writing proficiency, regardless of the instructor's experience with the teaching model. This does not, of course, mean that teachers do not have any influence on learning, and thus are not necessary in a flipped learning environment. On the contrary, it could be argued that the teachers are the most vital aspect of the learning process in the flipped classroom; they provide the necessary personalized coaching for students that is made more possible within such a learning environment. Teachers who are not experienced with the use of computer technology or the flipped learning approach may feel overwhelmed when incorporating this way of instruction in their classes. Based on the results of this study, and previous research, various pedagogical implications can be discussed for language instructors, both experienced and new to the flipped learning idea.

First, some of the videos used in this investigation seem to have been too long. Due to the limited concentration spans of students, Bergmann and Sams (2012) recommend that videos be no longer than 10 minutes, and in cases in which the videos are longer than this, they should be broken up into two or more shorter videos of around 5 minutes. Although the teacher who was creating the videos did attempt to keep them as brief as possible, some videos were longer, with the longest one being close to 15 minutes. This was recognized by at least one student in the feedback at the end of the course: "Sometimes there was too much information in the videos, so I want the teacher to consider this in the future." It is recommended that teachers new to the flipped learning approach make an extra effort to keep the information in the videos as concise as possible.

Second, it may be effective for the person giving the explanations and creating the videos for the flipped classroom to meet the students who are taking the lessons. While this may be difficult in cases, as with the present study, in which the creator lives a long distance from the students, a short greeting or introduction, for example by using online video chat sites, may increase the familiarity students feel with the explanations. Although similar increases in proficiency were seen for both groups in this study, there were times during class when the teacher of the NFE Group referred to the FE teacher (i.e., the teacher who created the videos). If the students of the NFE Group had had some personal interaction with the FE Group teacher, they might have felt more comfortable with the "semi-team-teaching" learning environment at an earlier stage.

Third, as suggested by Lee and Wallace (2017) and mentioned previously in this paper, the number of studies focusing on the flipped learning model is still remarkably low. There is still a need for investigations that compare increases in proficiency (or lack thereof) while conducting comparisons of the same learning content in a flipped classroom and a traditional one. As mentioned earlier in the paper, some studies have focused on this, but much more research is necessary to strengthen the support for flipped learning while also identifying possible weaknesses with the approach.

7. Conclusions

The results of this study suggest that the flipped learning approach is effective for improving students' English composition skills, regardless of the experience of the teacher. The study is, however, limited by a number of weaknesses.

First, the discussion of the results in this study is, admittedly, purely based upon the analyses of the data. Although the students were asked to voluntarily share their opinions about the flipped learning model, very few students, in fact, gave responses. Initially, it was thought by the authors that if students were required to give feedback, they might not give their true opinions, but just write what they felt the researchers

wanted to hear. In future studies, it will be beneficial to use a mixed-method approach by including, for example, interviews in the process of gathering data. Through the interviews, the researchers will be able to obtain qualitative data to support and strengthen the ideas discussed in this paper regarding agency and autonomy. Furthermore, interviewing students might give researchers a clearer answer to the chief question of this study of whether the amount of experience the teacher has with flipped learning influences the amount of success students have in their learning.

Second, an analysis of the compositions and the type of language being used throughout might have given the researchers an insight into improvements in the quality of the students' writing. Even though a rubric was provided for the markers of the compositions in this study in order to get reliable evaluations, it would be beneficial to also consider online analyses of the compositions to objectively test the content. Readability formulas, such as Gunning Fog and Flesch Reading Ease, may provide further independent indications of developments in students' writing quality as a result of studying in a flipped learning environment and whether these improvements, or lack thereof, differ depending on their teacher.

The flipped classroom approach has proven, in the majority of cases in which it has been tested, to be highly successful in improving the positivity with which students approach their learning. Although this has been investigated in many fields, especially science, mathematics, and medicine, there has been a gradual increase in the literature supporting the use of flipped learning in the field of second language acquisition. In the present study, the authors have attempted to provide further evidence as to whether the flipped learning approach brings about a significant increase in the EFL composition proficiency of university students. The results have indicated that it was successful, not only for students studying under a teacher who had ample experience of the flipped learning model but also for those whose teacher was using this approach to teaching for the first time. With further investigations concentrating on the effects of the flipped learning approach on students' proficiency, motivation, and attitudes to learning, more hints may become clear for language teachers to conduct their classes in the most effective and efficient ways possible.

Acknowledgements

The authors wish to thank the participating students for their contributions to this study.

References

- Anderson, J. (1986). *Taking charge: Responsibility for one's own learning*. Unpublished MA Thesis. The School for International Training, Brattleboro, VT.
- Baker, J. W. (2000). The "classroom flip": Using web course management tools to become the guide by the side. In *Selected Papers from the 11th International Conference on College Teaching and Learning*, 9-17.
- Benson, P. (2011). *Teaching and researching autonomy*. Harlow, England: Pearson Education.
- Bergmann, J., & Sams, A. (2012). *Flip your classroom: Reach every student in every class every day*. Eugene, OR: International Society for Technology in Education.
- Brinks Lockwood, R. (2014). *Flip it!: Strategies for the ESL classroom*. Detroit, MI: University of Michigan Press.
- Crouch, C. H., & Mazur, E. (2001). Peer instruction: Ten years of experience and results. *American Journal of Physics*, 69(9), 970-977. doi: 10.1119/1.1374249
- Dam, L., & Legenhausen, L. (1996). The acquisition of vocabulary in an autonomous learning environment-the first months of learning English. In R. Pemberton, E.S. Li Li, W.F. Or, & H.D. Pierson (eds). *Taking control: Autonomy in language learning*. Hong Kong: Hong Kong University Press, 265-280.

Forsythe, E. (2016). Pedagogical rationale for flipped learning and digital technology in second language acquisition. In Information Res Management Association (ed.), *Flipped instruction: Breakthroughs in research and practice*. Information Science Reference, 116-130. doi: [10.4018/978-1-5225-1803-7.ch007](https://doi.org/10.4018/978-1-5225-1803-7.ch007)

Fraser, C. A. (2007). Reading rate in L1 Mandarin Chinese and L2 English across five reading tasks. *The Modern Language Journal*, 91(3), 372-394. doi: [10.1111/j.1540-4781.2007.00587.x](https://doi.org/10.1111/j.1540-4781.2007.00587.x)

Harumi, S. (2011). Classroom silence: Voices from Japanese EFL learners. *ETL Journal*, 65, 260-269. doi: [10.1093/elt/ccq046](https://doi.org/10.1093/elt/ccq046)

Holec, H. (1981). *Autonomy in foreign language learning*. Strasbourg, France: Council of Europe.

Hung, H. T. (2015). Flipping the classroom for English language learners to foster active learning. *Computer Assisted Language Learning*, 28, 81-96. doi: [10.1080/09588221.2014.967701](https://doi.org/10.1080/09588221.2014.967701)

Keefe, J. W. (2007). What is personalization? *Phi Delta Kappan*, 89(3), 217-223. Retrieved from <http://journals.sagepub.com/doi/pdf/10.1177/003172170708900312>

Lage, M. J., Platt, G. J., & Treglia, M. (2000). Inverting the classroom: A gateway to creating an inclusive learning environment. *The Journal of Economic Education*, 31(1), 30-43. doi: [10.2307/1183338](https://doi.org/10.2307/1183338)

Larson-Hall, J. (2008). Weighing the benefits of studying a foreign language at a younger starting age in a minimal input situation. *Second Language Research*, 24(1), 35-63. doi: [10.1177/0267658307082981](https://doi.org/10.1177/0267658307082981)

Lee, G., & Wallace, A. (2017). Flipped learning in the English as a foreign language classroom: Outcomes and perceptions. *TESOL Q*, 1-23. doi: [10.1002/tesq.372](https://doi.org/10.1002/tesq.372)

Lee, P. J. (2017, June). *Effects of interactive subtitles on EFL learners' content comprehension and vocabulary learning*. Paper presented at JALTCALL2017 Conference, Matsuyama, Japan.

Leis, A. (2015). Dynamics of effort in flipped classrooms in an EFL environment. *Educational Informatics Research*, 14, 15-26. Retrieved from https://www.researchgate.net/publication/289530659_Dynamics_of_Effort_in_Flipped_Classrooms_in_an_EFL_Environment

Leis, A. (2016). Flipped learning and EFL proficiency: An empirical study. *Journal of the Tohoku English Language Education Society*, 36, 77-90. Retrieved from https://www.researchgate.net/publication/303759005_Flipped_Learning_and_EFL_Proficiency_An_Empirical_Study

Leis, A., Cooke, S., & Tohei, A. (2015). The effects of flipped classrooms on English composition writing in an EFL environment. *International Journal of Computer-Assisted Language Learning and Teaching*, 5(4), 37-51. doi: [10.4018/978-1-5225-0783-3.ch062](https://doi.org/10.4018/978-1-5225-0783-3.ch062)

Lim, K. M., & Hui Zhong, S. (2006). Integration of computers into an EFL reading classroom. *ReCALL*, 18(2), 212-229. doi: [10.1017/s0958344006000528](https://doi.org/10.1017/s0958344006000528)

Little, D. (1990). Autonomy in language learning. In I. Gathercole (ed.) *Autonomy in language learning*, London, England: CILT, 7-15.

Loewen, S., & Plonsky, L. (2016). *An A - Z of applied linguistics research methods*. New York, NY: Palgrave Macmillan.

Lyster, R. (2004). Differential effects of prompts and recasts in form-focused instruction. *Studies in Second Language Acquisition*, 26(4), 399-432. doi: [10.1017/s02722263104263021](https://doi.org/10.1017/s02722263104263021)

Matsukawa, R., & Tachibana, Y. (1996) Junior high school students' motivation towards English learning: A cross-national comparison between Japan and China. *ARELE: Annual Review of English Language Education in Japan*, 7, 49-58. Retrieved from https://www.jstage.jst.go.jp/article/arele/7/0/7_KJ00007108213/_pdf

Mazur, E. (1997). Peer instruction: Getting students to think in class. *AIP Conference Proceedings*, 981-988. doi: 10.1063/1.53199

Mehring, J. G. (2015). *An exploratory study of the lived experiences of Japanese undergraduate EFL students in the flipped classroom*. (Doctoral dissertation, Pepperdine University). Retrieved from https://www.academia.edu/12538458/An_Exploratory_Study_of_the_Lived_Experiences_of_Japanese_Undergraduate_EFL_Students_in_the_Flipped_Classroom

Perez, M. M., Norgate, W. V. D., & Desmet, P. (2013) Captioned video for L2 listening and vocabulary learning: A meta-analysis. *System*, 41(3), 720-739. doi: 10.1016/j.system.2013.07.013

Sasaki, M. (2011). Effects of various lengths of study-abroad experience on Japanese EFL students' L2 writing ability and motivation: A longitudinal study. *TESOL Quarterly*, 45(1), 81-105. doi: 10.5054/tq.2011.240861

Yang, J. C., & Chang, P. (2014) Captions and reduced forms instruction: The impact on EFL students' listening comprehension. *ReCALL*, 26(1), 44-61. doi: 10.1017/s0958344013000219

Research paper

Digital flashcard L2 Vocabulary learning out-performs traditional flashcards at lower proficiency levels: A mixed-methods study of 139 Japanese university students

Robert John Ashcroft*, Robert Cvitkovic* and Max Praver**
*Tokai University, Japan | **Meijo University, Japan

*bob.ashcroft@tsc.u-tokai.ac.jp | **bcvitkovic@yahoo.ca | *** praver@meijo-u.ac.jp

Abstract

This study investigates the effect of using digital flashcards on L2 vocabulary learning compared to using paper flashcards, at different levels of English proficiency. Although flashcards are generally believed to be one of the most efficient vocabulary study techniques available, little empirical data is available in terms of the comparative effectiveness of digital flashcards, and at different levels of student English proficiency. This study used a mixed-methods experimental design. The between-subjects factor was English Proficiency consisting of three groups: basic, intermediate and advanced. All participants underwent both a digital flashcards treatment and paper flashcards treatment using words from the Academic Words List. For each study mode, the two dependent variables were Immediate, and Delayed Relative Vocabulary Gain. The results of this study indicated that Japanese university students of lower levels of English proficiency have significantly higher vocabulary learning gains when using digital flashcards than when using paper flashcards. Students at higher levels of proficiency performed equally well using both study modes. It appears that by compensating for the gap in metacognitive awareness and effective learning strategies between students of lower and higher levels of language proficiency, digital flashcards may provide the additional support lower-level learners need to match their advanced-level peers in terms of their rate of deliberate vocabulary acquisition.

Keywords: Vocabulary, digital flashcards, paired-associates, autonomy, English proficiency, Academic Words List.

1. Introduction

The exponential growth and development of computer technology is having a significant impact on many aspects of foreign language pedagogy. Most teachers intuitively recognize the opportunities afforded by Computer Assisted Language Learning (CALL) materials and strive to integrate these technological innovations into their teaching practices. However, due to the rapid rate of change, related research and accompanying pedagogy can often lag behind the development of new CALL applications. As a result, teachers may lack the support of a theoretical framework. Moreover, digital technology often appears intrinsically desirable in itself, rather than because of the possible objective benefits to students. This superficial appeal, combined with a lack of pedagogy, can result in unrealistic expectations of CALL applications in terms of learning outcomes (Gartner, 2017). It is difficult for language teachers to recognize which CALL applications will enhance student learning and which will not. Although the influence of CALL is felt throughout most aspects of language teaching and learning, the present study specifically examines vocabulary learning.

One way in which CALL technology has influenced the field of L2 vocabulary learning has been with the emergence of digital flashcards, with Quizlet and Anki being popular examples. Despite the widespread use of these applications among students and teachers, the comparative effectiveness of digitized flashcards remains under-

researched (Nation and Webb, 2011). The authors of the present study used the Quizlet application in an effort to determine the effectiveness of digital flashcards for learning L2 vocabulary compared to the traditional paper variety at different student proficiency levels. The results of the experiment are summarized in the five points below.

- Immediate vocabulary gains are significantly higher for digital flashcards than paper flashcards for both intermediate and lower levels of English proficiency.
- There is no significant difference between digital flashcards and paper flashcards for immediate vocabulary gains for higher English proficiency levels.
- Delayed vocabulary gains are significantly higher for digital flashcards than paper flashcards for students of lower levels of English proficiency.
- There is no significant difference between digital flashcards and paper flashcards for delayed vocabulary gains for intermediate English proficiency levels.
- Delayed vocabulary gains are significantly higher for paper flashcards than digital flashcards at higher levels of English proficiency.

2. Literature review

The following sections describe the relative merits of using paper and digital flashcards for vocabulary learning. The description includes a detailed consideration of how digital flashcards might further enhance the benefits inherent in paper flashcards. This is followed by a summary of a number of studies into the relative effectiveness of CALL and traditional vocabulary learning.

2.1. Paper flashcards

Research suggests that using paper flashcards is one of the most efficient means of deliberate vocabulary study techniques available (Elgort, 2010). Also known as paired-associate learning, this technique involves using small cards with the target L2 word on one side and the meaning of that word on the other. Using flashcards is thought to be particularly effective due to a combination of factors. First, because flashcards are portable, and therefore convenient, they can help engender student autonomy (Nation, 1995, 2003, 2005). The freedom to study whenever and wherever they like can have a liberating effect on students. Second, flashcards facilitate spaced-learning (Nation, 2003) where students can revisit items over an extended period (Hulstijn, 2001; Webb, 2007). The positive effect of spaced-learning on vocabulary acquisition is thought to be particularly strong. A further advantage of flashcards is that they can be grouped into sets (Cohen, 1990) based on relevant criteria such as lexical groups or test items. Finally, flashcards can include L1 translations, providing a visual link between L1 and the target language (Cross & James, 2001) and thereby further adding to their positive motivational effect. Although the meaning of the target word can be conveyed in several ways, such as a picture, L2 definition or L2 synonym, research indicates that L1 translation is the most effective method (Laufer & Shmueli, 1997; Nation & Webb, 2011). Using flashcards is more convenient, allows spaced-learning, and can include L1 synonyms of target vocabulary.

There are several points to bear in mind when trying to optimise the effectiveness of vocabulary flashcards. Baddeley (1990) stresses the importance of the retrieval process. Once a new word and its L1 meaning have been met, at the next meeting the student should see only the target word and try to recall the L1 meaning. It is also important to continually change the order of the flashcards (Nation & Webb, 2011). This prevents previous items from triggering the memory of subsequent ones, and also allows students to focus on the more difficult items. The optimal number of items for a study set is also an important consideration. Suppes and Crothers (1967) found that for lower level students it should be around 20 cards, and for more advanced learners, up to 50 cards is acceptable. Using paper flashcards to learn vocabulary is a long-standing technique and has been thoroughly researched. However, a more recent development is the emergence of computer-based, digital flashcard applications.

2.2. Digital flashcards

Several Web 2.0 flashcard applications now allow users to create, study and share with digital flashcards. The digital flashcard application chosen as the focus of this experiment was Quizlet, a popular choice for many students and teachers. The site has more than 20 million monthly users and over 140 million freely available user-made flashcard sets (Quizlet, 2017). The application has an attractive, intuitive interface and requires little set up or computer know-how to start studying new words. The website allows teachers to create a virtual class, and invite students to join. Once students have joined a virtual class they have access to the study sets within the class and can track their progress and that of other members of the class. Teachers have access to information about the study behavior and performance of the class members. There is also a Quizlet app available to download to a mobile device both for Android and i-OS. The app allows flashcard sets to be downloaded to the device and used with or without Internet connection.

The Substitution Augmentation Modification Redefinition (SAMR) Model (Puentedura, 2012) provides a means of assessing the integration of technology and its effect on teaching and learning. It divides CALL innovation into four stages of progressively greater degrees of enhancement. A study by Ashcroft and Imrie (2014) used the SAMR Model to assess the impact of using Quizlet vocabulary flashcards compared to paper vocabulary flashcards. They concluded that digital flashcards might be more effective due to additional features such as audio, immediate feedback, a seamless and user-friendly interface, and their high accessibility through a range of platforms. This additional functionality of Quizlet for L2 vocabulary study, compared to traditional flashcards is discussed in detail below using a framework adapted from Reinders and White (2011) outlining areas in which CALL materials in general can have pedagogical advantages over traditional teaching materials.

2.2.1. New activities

New types of activities are made possible using CALL applications which would be difficult or even impossible with traditional teaching materials. Indeed, this can be said of the Quizlet website, which offers a choice of four study modes. Firstly, the Flashcard Mode includes automated audio rendition of words on the cards. Next, the Learn Mode presents one side of a flashcard and requires the hidden item on the reverse of the card to be entered using the keyboard. If the target word is typed correctly, the program moves on to display the next card. If not, the answer is given, and learners are required to retype the word. The third study mode is Test, where users can set test parameters, such as the number and type (multiple choice, written, true/false, or matching) of questions. The test is generated based on the parameters and users are required to complete the test using the keyboard and mouse. When the test is complete, the total score is displayed, along with a list of the test items including the students' responses, the correct answer, and whether students answered each item correctly or not. The last study mode is called Spell. Here users must enter the target item using the keyboard based on an audio prompt.

In addition to the study modes, there are also two game modes. The first game is receptive. The user must match paired associates against the clock. The app then challenges users to beat their best time. The other game is a productive activity where users must enter the target item when prompted by one half of a paired associate. This must be done before an asteroid crashes into the planet below. Points are awarded for pairs successfully matched. Asteroids fall progressively faster, increasing the difficulty of the activity the further users progress. There is clearly a much greater variety of activities available to vocabulary learners using Quizlet compared to paper flashcards.

2.2.2. Feedback

Immediate feedback dependent on users' input is possible when using CALL materials. The Quizlet app provides high density feedback to users. For example, in the Learn study mode, the app will signal whether an item has been entered correctly or not. If users type in the answer for a different card, there is a confusion alert. A message appears explaining the problem and displaying the card correctly matching the entered response. When learners have worked through all the cards in a study set, all items are

then displayed, indicating whether each one was answered correctly or not. A percentage correct total is also shown. Learners are then required to repeat the process for those items answered incorrectly in the previous round. This process is repeated until the correct total reaches 100 percent. The Quizlet app allows for far richer and more immediate feedback than using paper flashcards.

2.2.3. Non-linearity

Traditional foreign language classrooms typically progress in lockstep (Richards & Schmidt, 2002) fashion with all students transitioning together from one activity to the next under the supervision of the teacher. However, CALL materials offer individual students many more study choices and the freedom to use these in any order and for as long as they wish. This holds true for Quizlet, with four study modes and two game formats. Students can approach learning by using any of the modes and in any order they choose. Moreover, in all modes, Quizlet presents cards to the user in randomized order, an important factor that maximizes vocabulary learning (Nation & Webb, 2011). In addition, flashcard sets can be combined, and individual flashcards can be starred to focus on more challenging words.

2.2.4. Monitoring and recording progress

Many CALL applications have the capability of monitoring and recording progress. This information can be made available to teachers, allowing those students not doing the work to be identified, if necessary. Monitoring data can also be used by the application itself to modify future studying activity. If a student is a member of a Quizlet virtual class, the website tracks the user's performance, and this data is available to students and teachers. The site allows flashcards to be sorted according to users' past performance. The cards are displayed in order from those most to those least often missed. This provides the opportunity for students to reflect on the learning process, and to target more problematic vocabulary. In addition, in the game mode Asteroids, Quizlet recycles words which students have missed earlier in the game.

2.2.5. Control

CALL materials provide a greater degree of control for students than traditional materials. Quizlet allows the option of studying with a subset of cards which have been missed by learners in previous study sessions. This ability to target words based on individualized feedback provides a greater sense of control over the learning process for students. The availability of Quizlet on a personal computer, tablet, or smart phone also passes greater control to users. Increased levels of control provide opportunities for the development of metacognitive skills and learner autonomy (Reinders and White, 2011). The Quizlet application provides additional functionality and control previously unavailable through traditional analogue flashcard use.

2.3 Paper versus digital flashcards: existing research

Although the benefits of the additional functionality of Quizlet over paper flashcards seem apparent, the results from existing empirical studies which examine the relative effectiveness of digital flashcards over their traditional counterparts remain inconclusive. One study of 226 Japanese high school students examined the comparative effect on vocabulary gains of using word lists, word cards and a CALL application to study ten vocabulary items (Nakata, 2008). The experiment found no significant difference between using paper flashcards and the computer application. A further study by Lees (2013) also found no difference between the effectiveness of paper flashcards and Quizlet flashcards. Another study by Hirschel and Fritz (2013) used CALL-based vocabulary learning and vocabulary notebooks with 140 university students in Japan. The results showed no significant difference between the two study modes. A further study also found no significant difference between using paper flashcards and internet based digital cards, however the results did show a significant difference between paper flashcards and digital flashcards available on a mobile device, such as a smart phone or tablet (Nikoopour, Jahanbakhsh & Azin, 2014). In all these studies, participants included only those from within the same English proficiency level band, or mixed level homogenized groups, so none of the results could take into account possible effects of proficiency level on the relative effectiveness of the

treatments. The present study attempts to address this gap in the research by answering the following research question:

RQ1. Does student English proficiency level influence the relative effectiveness of digital and paper flashcards in terms of L2 vocabulary learning gains?

3. Method

The purpose of this study was to investigate any difference in effect of using digital flashcards compared to paper flashcards, and to determine whether students English proficiency level also influenced the effectiveness of either study mode. Participants underwent both digital and paper flashcard treatments. For each treatment, a pre-test was used to determine how many of the target words were already known to each participant, and a post-test indicated how many items had been learned due to the treatment. A delayed post-test measured the rate of attrition of this learning. Details of how the experiment was carried out are provided in the subsections which follow.

3.1. Participants

The participants were 139 native Japanese, English language undergraduate students at a large university in Japan. Ages ranged from 18 to 24 years old, with 64 male and 75 female participants. All participants had received formal English instruction for at least seven years. Students belonged to either basic ($n = 32$), intermediate ($n = 46$) or advanced ($n = 61$) level integrated skills-based English classes. Students had been placed at either basic, intermediate or advanced-level based on a university administered TOEIC listening and reading test, compulsorily taken by all students at the start of their freshman year. A total of seven classes participated in the research: two basic-level classes, two intermediate, and three advanced-level classes. All participants owned a smart phone (either iPhone or Android). Many of the students participating in this research were also enrolled in different English courses through the university during the experimental period. The TOEIC listening and reading score ranges are shown in Table 1.

Class Level	Class #	Class n	TOEIC (L&R)	Level n
Basic	1	15	under 230	32
	2	17		
Intermediate	3	21	230 to 550	46
	4	25		
Advanced	5	21	over 550	61
	6	17		
	7	23		

3.2. Design

This study used a mixed-methods experimental design. The within-subjects factor was Study Mode, of which there were two levels, digital flashcards and paper flashcards. The between-subjects factor was English Proficiency, which consisted of three levels, basic, intermediate and advanced. The two dependent variables were Immediate and Delayed Relative Vocabulary Gain. These were defined as the number of new words learned from a closed set of target words expressed as a proportion of those words unknown prior to treatment. Word knowledge was measured using a productive L2 measure which was

prompted with the L1 half of a paired associate, along with the first letter of the L2 target word. A productive measure of vocabulary was chosen because many of the students taking part in the research were also taking English academic writing classes. The authors concluded that a productive measure of vocabulary gains was more appropriately matched to the current academic needs of the students.

3.3. Target vocabulary

A fixed and relatively small set of words was used for several reasons. Firstly, this helped the experiment to reflect the targeted nature of vocabulary study using flashcards in real-world learning contexts, thereby increasing the ecological validity of the design. In addition, measuring learning gains using achievement pre- and post-tests would allow a more precise measure of progress. Measuring changes in overall vocabulary size would, in contrast, be much more problematic since incremental gains would be proportionally very small. A further reason for the use of a small number of specific items was that the treatment period could be kept comparatively short, thereby minimizing the probability of students meeting target words outside the treatments.

The selection of words was informed according to two criteria. Firstly, it was important that items should be largely unknown to the participants so as to allow the effect of the treatments to be measured. Secondly, to maintain ecological validity, it was important that the items were relevant and useful for students to know. Using words from the Academic Words List (AWL) (Coxhead, 2000) allowed both requirements to be accommodated. The AWL was created through the analysis of a corpus of around 3,500,000 running words of written academic text. The list contains 570 word families, divided into nine sub-lists of 60 and a tenth sub-list of 30. Word families are ordered from the most frequent (List 1) to the least frequent (List 10). Coxhead used the most frequent form of each word family, per the academic corpus, when compiling the AWL.

A total of 120 words were used in the present study. For both treatments (digital and paper flashcards), participants studied with a different set of 60 target words. AWL Sub-lists 1 and 2 (representing 120 word families) were selected for this purpose as they are the most frequent AWL words and therefore most likely to be useful for students to know. Using computer randomization software, thirty words were selected at random from List 1, and then combined with 30 randomly selected words from list 2. These 60 words constituted the paper flashcards study set. The remaining 60 words were used for the digital flashcards study set. Randomly selecting words in this way ensured that the two study sets had comparable mean frequencies in the academic corpus, and thereby removed any distorting effect of frequency differentials across study sets.

3.4. Instrumentation

Six dependent measures were administered to each participant. Pre-post, non-identical, 30-item vocabulary tests were developed to measure vocabulary gains for each treatment. For the pre-tests, test items were created for 30 words selected at random from each study set (60 total words). The remaining 30 words in each study set were used for the post-test items. Pre- and post-test scores (each out of 30) were used to calculate the relative vocabulary gains from the study set as a whole. The delayed-post tests were made by selecting 15 items at random from the corresponding pre- and post-tests.

The prompts on the pre- and post-tests included the Japanese equivalent of the target item, as well as a sentence in English with the target word omitted (see Table 2). The first letter of the target word was provided to discourage participants from producing synonyms for target answers (Hughes, 2013). The tests were administered in paper format to allow flexibility for misspelt answers, and for British / American English variations. As a rule, two letters per item were permitted to be misspelt. Both American and British spellings of answers were accepted. The L2 words from the flashcards of the respective treatment were the only acceptable answers.

L1 prompt	prompt	answer
財政	He studied banking and (f_____) at business school.	finance

Using different items for pre- and post-tests ensured that learning of target vocabulary due to taking the pre-test and the consequent distorting effect on the results of the post-test could be avoided. However, because the items on the pre-test differed from those on the corresponding post-test, and delayed post-test, it was necessary to check that the tests were of equivalent difficulty. In order to do this, the two sets of pre-, post-, and delayed post-tests were administered to 176 native Japanese students of English studying at the same university as those in the main experiment. The test validation group received no treatment. Like the experimental procedure, the digital cards pre-test was administered in class 1, and the post-test in class 3. The paper cards pre-test was done in class 4 and the post-test in class 6. All participants therefore took all six tests. During the intervening class time, students did not study any vocabulary from any of the tests. Paired-samples *t*-tests were conducted between each of the six combinations of tests. The *t*-tests included a Bonferroni correction to offset for the number of tests. No significant difference in the mean scores for any combination of the tests was found (see Table 3). This result indicated that the tests were of equal difficulty and that any difference in the test scores in the main experiment should be attributed to the intervening treatment.

Combination	<i>t</i> (350)	<i>p</i>	<i>M</i>	<i>SD</i>
QPre/QPost	-0.543 a	0.59*	-0.07	1.81
QPre/PPre	0.000 b	1.00*	0.00	1.62
QPre/PPost	.208 b	0.84*	0.03	1.81
QPost/PPre	.334 b	0.74*	0.07	2.94
QPost/PPost	.429 b	0.67*	0.10	3.17
PPre/PPost	.131 b	0.89*	0.03	2.88

a: df=75; b: df=174; *:p>.05

3.5. Treatments

The research was conducted over the course of two semesters. The experimental design included two treatment cycles: a paper flashcard treatment and a digital flashcard treatment. Each treatment cycle spanned three 90-minute classes, making a total of six classes for both treatments combined (see Table 4). This was a repeated measures design, with all participants receiving both treatments. However, the order of treatments was varied. Participants were arbitrarily divided into two groups: The Digital First group (*n*=75: 17 basic, 25 intermediate, and 23 advanced) received the digital flashcard treatment first, followed by paper flashcards. The Paper First" group (*n*=74: 15 basic, 21 intermediate, and 38 advanced) worked with paper flashcards first,

followed by digital flashcards. Adopting this counter-balanced design helped to minimize the effect of order of study mode on the results.

Table 4. Repeated measures experimental design protocol		
Class #	Digital First Group	Paper First Group
1	Digital Flashcards Pre-test Quizlet Orientation	Paper Flashcards Pre-test; Writing out Flashcards
2	Digital Treatment Homework assigned	Paper Treatment; Homework assigned
3	Digital Flashcards Post-test	Paper Flashcards Post-test
4	Paper Flashcards Pre-test Writing out Flashcards	Digital Flashcards Pre-test Quizlet Orientation
5	Paper Treatment Homework assigned	Digital Treatment Homework assigned
6	Paper Flashcards Post-test Digital Flashcards Delayed Post-test	Digital Flashcards Post-test Paper Flashcards Delayed Post-test
9	Paper Flashcards Delayed Post-test	Digital Flashcards Delayed Post-test

Each flashcard, both paper and digital, had a target English item on one side and the corresponding Japanese (L1) translation of the most frequent academic meaning according to the Oxford English dictionary (Soanes, 2010) on the other. The English was translated into Japanese by a native Japanese university English teacher. The translations were then confirmed by a second Japanese native of high English proficiency. The prompts on the pre- and post-tests included the same Japanese equivalent of the target item (see Table 2). For both treatments, there was a pre-test, a post-test after the treatment, and a delayed post-test 3 weeks after the treatment. The post-test was included in an effort to measure comparative differences in the rates of attrition of vocabulary learning with the two learning modes, and at different levels of proficiency. A three-week gap between immediate post and delayed post-tests was chosen to reflect what might be a realistic period for the participants between deliberate vocabulary study and the opportunity to use these words again incidentally in the course of their studies.

3.5.1. Digital treatment

The digital flashcard treatment sessions were conducted in a CALL classroom. Students used the Quizlet website, not the downloadable app version. Each student had the use of a computer with Internet connection and a pair of headphones. Additionally, each pair of students could see a monitor showing the teacher's computer screen display. The digital flashcards were prepared on the Quizlet site in advance of the treatment. The cards were organized into three sets of 20 items. In the first 90-minute class of the treatment, the instructor introduced students to the Quizlet application. Students registered with the Quizlet website, and then joined a virtual class (see section *Digital Flashcards*), created by the researchers in advance. Using a prepared set of 20 flashcards based on items taken from the AWL List 3 (i.e., not the items being used for this research), students practiced using the site. These flashcards were prepared with the English item on one side, and the Japanese equivalent on the other. The class was conducted in lockstep (Richards & Schmidt, 2002). Students first examined the flashcards and then were shown how to operate each of the various study modes using the shared monitors. The teacher also demonstrated the audio function available in some of the study modes, which the students then went on to use. During this stage,

the teacher periodically displayed summaries of class progress on the shared monitors. For the last 20 minutes of class, students were shown how to download the Quizlet app to their smart phone. There was no homework assigned, and the deck of 20 items was removed from the virtual class after this session.

The next class began with the pre-test for the digital flashcard word set. Then the students studied with digital flashcards for the remaining 90 minutes using the three sets of 20 items from the digital flashcard study set. Students were free to use any study mode and sets of digital cards in any order they wished. The teacher monitored participants closely, offering assistance and ensuring they remained on task. At the end of the session, students were told how to download the digital flashcard study sets to their mobile phones. For homework, students were told to study the digital flashcards study set with Quizlet using a PC or on their smart phones in preparation for a test to be given next class. No attempt was made to measure Quizlet usage outside class. In the next session, students took the digital flashcard post-test. The digital flashcards delayed post-test was given exactly three weeks after the digital flashcards immediate post-test was administered.

3.5.2. Paper treatment

All paper flashcard treatment sessions were conducted in a standard classroom with whiteboard and moveable chairs and desks. Students were not permitted to use their cell phones. Firstly, the pre-test for the paper flashcards words set was administered. Students were then each given a set of 100 blank paper flashcards measuring 5cm by 2.5cm, bound by a plastic ring. The ring could be detached enabling the cards to be separated. The students were given a copy of the paper flashcard study set. This was on a sheet of A4 paper depicting a table with the 60 target vocabulary items in English in the first column and the corresponding Japanese equivalents in the adjacent column. Students copied the vocabulary onto their set of flashcards. Each card had the English word on one side and the Japanese equivalent on the other. Students then wrote their names on the top cover card of their set and the teacher collected all the flashcard sets and vocabulary tables.

In the following session, students were handed back their paper vocabulary card sets which they had made during the previous class. Students removed the plastic clip from their set of cards, and worked individually to separate the items into two sub-sets: the words they felt they understood, and the words they did not. This task allowed them to focus their efforts on those words which were unfamiliar to them. The students were given 20 minutes to memorize these words. For the next 20 minutes, students worked with a partner to test each other on the words. Pairs took turns to prompt each other with the English (L2) items to elicit the Japanese (L1) meaning. Having the students retrieve the target L1 item from memory using the L2 paired associate is thought to optimize vocabulary learning (Baddeley, 1990). Finally, students were arranged into small groups of four or five members. Students took turns prompting the other students in the group with L2. The first group member to give the correct L1 meaning for each item was awarded the corresponding card. The winner of each round was the student who had received the most cards. This stage lasted for 40 minutes. At the end of class, students were told to use their set of 60 paper flashcards to study in preparation for a test on the vocabulary next class. No attempt was made to measure students' flashcard usage outside class. At the beginning of the next class, the paper flashcard post-test was administered. The paper flashcards delayed post-test was given exactly three weeks after the paper flashcards immediate post-test.

3.6. *Calculating vocabulary gains*

Table 5 shows the mean pre-test (paper and digital tests) scores for each proficiency group.

Table 5. Mean pre-test scores by proficiency level

	basic	intermediate	advanced
Mean pre-test scores	1.42	3.72	8.12

The mean pre-test scores ($t(212) = 7.71, p=.00$) were significantly higher for the advanced level group, ($M=8.12, SD=4.78$), than for intermediate level group, ($M = 3.72, SD = 3.81$), who in turn had significantly higher, ($t(154)=4.53, p=.00$), pre-test scores than the basic level students, ($M = 1.42, SD=1.69$).

Thus, there was more room for improvement for basic students, compared to intermediate and advanced, and for intermediate participants compared to the advanced. Using the raw scores to measure vocabulary gains would lead to inflated gain scores at lower levels of proficiency. To correct for this, relative gain scores (Horst, Cobb & Meara, 1998) were used. This measure considers individual differences in starting positions and is calculated using the following formula:

$$\frac{\text{Post-test score} - \text{Pre-test score}}{\text{Highest possible score (30)} - \text{Pre-test score}}$$

The maximum possible relative gain score is 1.0. This represents when a participant gets the maximum score (30 for this study) in the post-test, irrespective of the pre-test result. Two relative gain scores were ascertained for each participant, one for the paper flashcards treatment, and the other for digital flashcards.

4. Results

Two, two-way mixed analyses of variance (ANOVA) were conducted to evaluate the effect of English proficiency and Study Mode on vocabulary gains. The dependent variables were Immediate Vocabulary Gain and Delayed Vocabulary Gain both measured from 0.00 to 1.00. The within-subjects factor was Study Mode, with two levels (Quizlet and paper flashcards). The between-subjects factor was English Proficiency with three levels (Basic, Intermediate, and Advanced).

4.1. Within-subjects main effect: study mode

For the sample as a whole, there was a significant main effect of Study Mode on Immediate Relative Vocabulary Gain, $F(1,136) = 12.87, p=.00, \eta^2 = .09$. Digital flashcard Immediate Gains ($M = .57, SD = .23$) were significantly higher than for paper flashcards ($M = .51, SD = .28$).

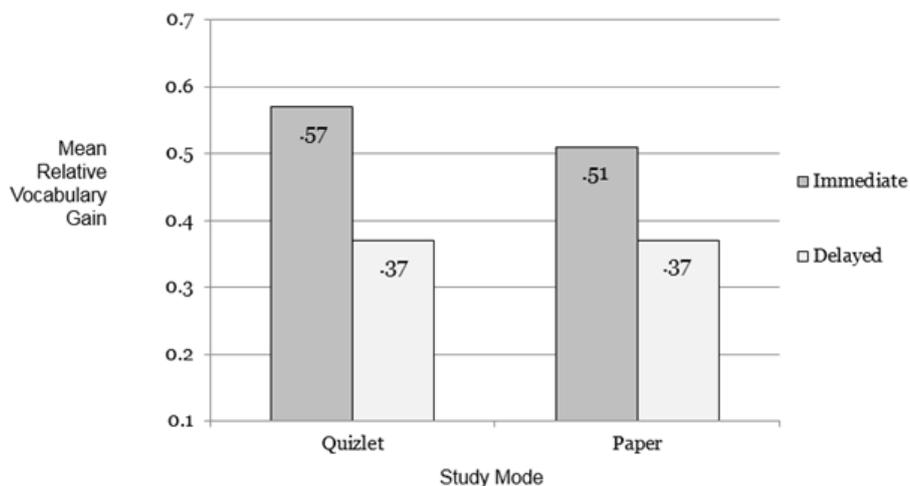


Figure 1. Immediate and Delayed Relative Mean Vocabulary Gains for Quizlet and Paper Flashcards.

However, there was no significant main effect of Study Mode on Delayed Relative Vocabulary Gain. Digital flashcard Delayed Gains ($M = .37, SD = .23$) were not

significantly higher than for paper flashcards ($M = .37, SD = .27$). This is can be seen in Figure 1.

4.2. Between-subjects main effect: English proficiency

There was a significant main effect of English Proficiency on Immediate Vocabulary Gain, $F(2,136) = 26.48, p=.00, \eta^2=.28$. Averaging for Study Mode, the advanced group scored significantly higher immediate gains ($M= .66, SD= .18$) than the intermediate group ($M= .50, SD= .22$), $t(105) = 4.28, p= .00$, who in turn did significantly better than the basic group ($M=.37, SD= .03$), $t(76) = 2.86, p=.005$.

There was also a significant main effect of English Proficiency on Delayed Vocabulary Gain, $F(2,136) = 30.61, p=.00, \eta^2=.31$. Averaging for Study Mode, the advanced group scored significantly higher delayed gains ($M= .48, SD= .19$) than the intermediate group ($M= .36, SD= .20$), $t(105) = 3.28, p=.00$, who in turn did significantly better than the basic group ($M= .17, SD= .13$), $t(76) = 4.53, p= .00$. These results can be seen in Figure 2.

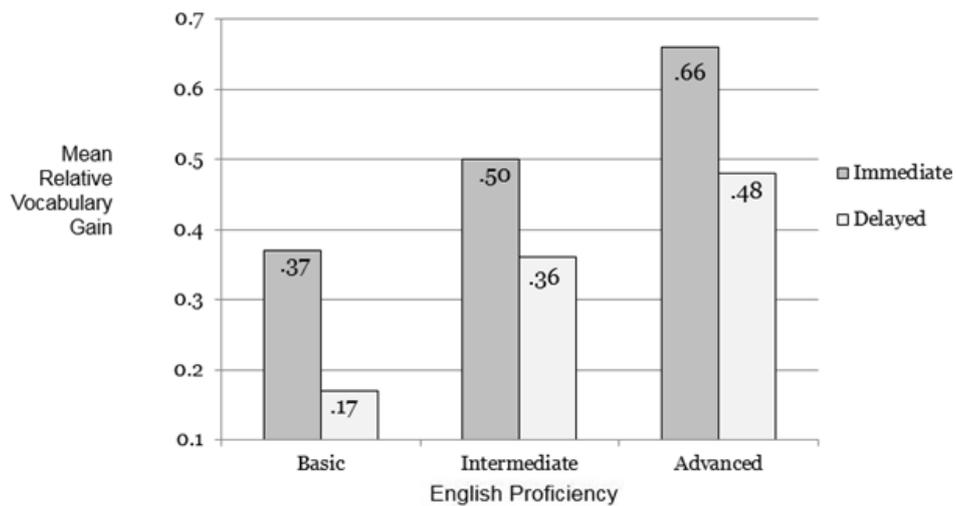


Figure 2. Immediate and Delayed Relative Mean Vocabulary Gains at different levels of English Proficiency.

4.3. Interaction effect: study mode and English proficiency

There was a significant interaction effect between the Study Mode and English Proficiency both for Immediate Relative Vocabulary Gains, $F(2,136) = 4.72, p=.01, \eta^2=.065$, and for Delayed Gains $F(2,136) = 8.42, p<.00, \eta^2=.11$. This indicates that the Immediate and Delayed Relative Vocabulary Gain scores differed according to Study Mode and also the level of English Proficiency. In other words, using digital or paper flashcards affected each proficiency level differently. To break this down, multiple comparisons were calculated at each level of English ability for both immediate and delayed gains.

Basic level participants achieved significantly higher Immediate Vocabulary Gains using Quizlet ($M= .43, SD= .21$) than when using paper flashcards ($M= .30, SD= .19$), $t(136) = 5.85, p= .00$. Intermediate level participants also, on average, achieved significantly higher immediate gains using Quizlet ($M= .55, SD= .22$) than when using paper flashcards ($M= .45, SD= .27$), $t(136) = 2.9, p=.00$. However, the immediate gain results suggest that there was no significant difference between using digital ($M= .66, SD=.21$) or paper flashcards ($M= .67, SD= .24$), $t(136)=-2.45, p=.81$, for the advanced level group. The relative immediate gains for each study mode and at each proficiency level are shown in Figure 3.

Using paper flashcards, the advanced group had significantly higher immediate vocabulary gains ($M= .67, SD= .24$) than the intermediate group ($M= .45, SD= .27$), $t(105)=4.38, p=.00$, and the intermediate group had significantly higher immediate gains ($M= .45, SD= .27$) than the basic group, ($M= .30, SD= .19$) $t(76)=2.72, p=.00$. Using digital flashcards, the advanced group had significantly

higher immediate vocabulary gains ($M= .66, SD= .21$) than the intermediate group ($M= .55, SD= .23$) $t(105)=2.573, p=.01$. The intermediate group had significantly higher gains ($M= .55, SD= .23$) than the basic group, ($M= .43, SD= .21$) $t(76)=2.390, p=.02$.

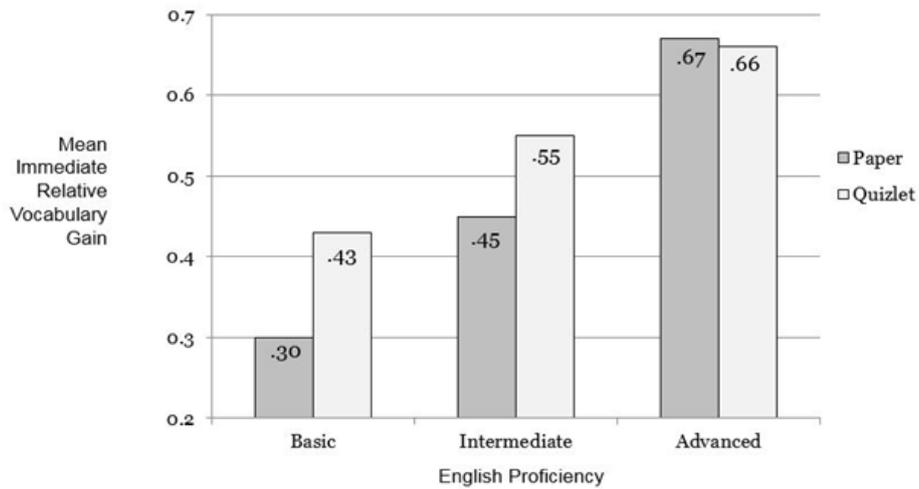


Figure 3. Immediate Relative Vocabulary Gains for Paper and Digital Flashcards for Different Levels of English Proficiency.

The delayed gain scores were significantly higher for basic students for digital flashcards ($M= .22, SD= .16$) than for paper flashcards ($M= .13, SD= .13$), $t(31) = 3.69, p = .001$. The delayed scores showed no significant difference between using digital or paper flashcards for the intermediate level group. Interestingly, for the advanced group the data showed that participants had significantly higher gain scores with paper flashcards ($M= .53, SD= .23$) than with digital flashcards ($M= .43, SD= .21$), $t(60) = 2.97, p = .004$. The relative delayed gains for each study mode and at each proficiency level are shown in Figure 4.

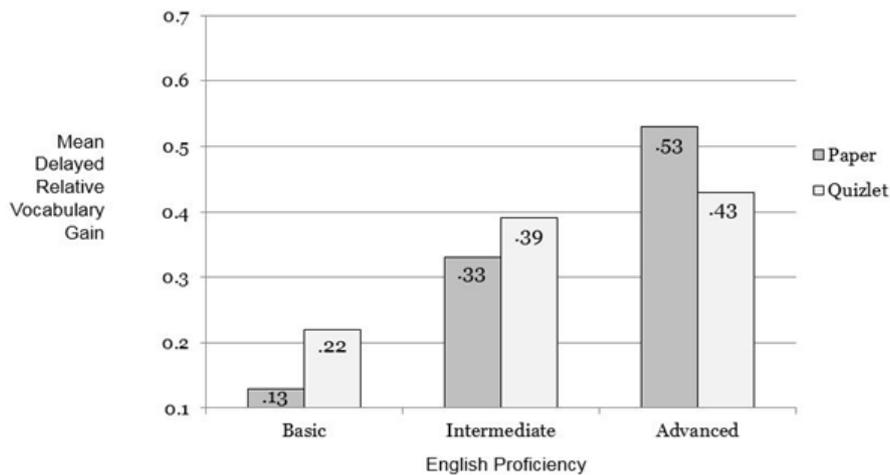


Figure 4. Delayed Relative Vocabulary Gains for Paper and Digital Flashcards for Different Levels of English Proficiency.

Using paper flashcards, the advanced group also had significantly higher delayed vocabulary gains ($M= .53, SD= .23$) than the intermediate group ($M= .33, SD= .24$) $t(105)=4.415, p=.00$, and the intermediate group had significantly higher delayed gains ($M= .33, SD= .24$) than the basic group, ($M= .13, SD= .31$) $t(76)=4.254, p=.00$. Using digital flashcards, the intermediate group had significantly higher delayed gains ($M= .39, SD= .25$) than the basic group, ($M= .22, SD= .16$) $t(76)=3.46, p=.001$. However, there was no significant difference between the advanced group and the intermediate group for delayed gains using digital flashcards.

5. Discussion and conclusions

The results show a significant main effect of proficiency level on both immediate and delayed vocabulary gain scores, indicating that student's proficiency level positively influenced their ability to learn new words, regardless of study mode. The effect of level on relative vocabulary gains is a striking observation. Unlike grammar, it is thought that vocabulary can be learned in any order, irrespective of proficiency level (Lightbrown & Spada, 1999). It would therefore be natural to assume that the words themselves would not account for such a proficiency level effect. A probable explanation is that those participants with better developed aptitudes to learning and higher levels of metacognitive awareness, have become more proficient in English as a result of these qualities. The English proficiency level groupings would therefore also correspond to increasing levels of metacognitive awareness and learner strategy development. Perhaps higher levels of focus, discipline, time management, motivation, confidence, and ability to apply learning strategies, which helped students become more proficient in English, also helped the same students to achieve greater vocabulary gains in this experiment, irrespective of whether they used digital or paper flashcards.

The analysis indicates that digital flashcards were more effective at increasing immediate vocabulary gains than paper flashcards for basic and intermediate-level students. Study mode had no significant effect on immediate vocabulary gains for advanced-level students. This suggests a negative correlation between proficiency, and the comparative superiority of digital flashcards over paper flashcards for L1 – L2 paired associate vocabulary learning. A possible explanation for this result might be that at lower levels the digital flashcards somehow compensated for the lack of metacognitive awareness and learner strategies. Perhaps certain characteristics of the digital application bolstered lower-level participants' lack of such qualities. The variety of activities offered by the Quizlet app, along with the high level of immediate feedback may have helped to boost and sustain the engagement and motivation of lower level students, which paper flashcards could not. The higher levels of control over their study provided by Quizlet due to the delineated nature of the app and access across multiple platforms may also have contributed to maintaining engagement and motivation at lower levels. It seems that by acting as a form of environmental support, the digital flashcards allowed lower level participants to perform to a higher level of proficiency. In contrast, in the experiment, the advanced level students achieved superior results irrespective of the study mode used. It seems that, at least for immediate gains, advanced level students may not require the features of digital flashcards in order to perform well.

The results for delayed vocabulary gains were, however, somewhat different. Again, basic students did significantly better using Quizlet than with the paper flashcards. Unlike the results for immediate gains, there was no significant difference in delayed gains between Quizlet and paper flashcards for the intermediate group. Moreover, the delayed gains for the advanced students were significantly lower using Quizlet than for paper flashcards. The results suggest that the negative correlation between proficiency and the superior effect of Quizlet over paper flashcards is even more pronounced for the delayed gains. In fact, the digital gains were lost significantly more quickly than the paper gains for the advanced level participants.

The higher rate of attrition of the vocabulary gains using Quizlet compared to those of paper flashcards could be attributable to a number of factors. It is possible that the advanced level group did not continue to study using the Quizlet flashcards after the immediate post-tests, while the intermediate and especially the basic group did. Another explanation is that the quality of learning was somehow different between proficiency levels using the Quizlet study cards. In order to discover why advanced gains were more susceptible to attrition, it would be useful to investigate how students used the study modes in the intervening period between the immediate and delayed post-tests. In addition, further research which examines the study behaviour of students outside class and their attitudes towards the two study modes would be helpful to find out how digital flashcards helped lower levels achieve higher learning gains than with paper. Qualitative data collected from paper and digital flashcard users at different

proficiency levels would help to shed light on how attitudes towards the two study modes differs according to level.

The results of this study suggest that digital flashcards help students at lower levels to achieve higher vocabulary gains than when they use paper flashcards. The most advanced group of students in this study however, did equally well with digital and paper flashcards. It seems that the extra functionality provided by the digital platform somehow compensated for lower-level participants' inability to study as effectively as advanced students when using paper flashcards. It seems plausible that lower levels of metacognitive awareness and effective learning strategies associated with lower proficiency level students was cancelled out when using the digital flashcards. This in turn may be due to features of digital flashcards such as greater variety of activities, high level of immediate feedback, increased sense of control and learner autonomy, and the non-linearity of the application. On the basis of these findings, it may be advisable for curriculum designers to consider including digital platforms for L2 vocabulary study for language learners at lower levels of proficiency.

References

- Ashcroft, R. J., & Imrie, A. C. (2014). Learning vocabulary with digital flashcards. *JALT2013 Conference Proceedings*, 639-646. Retrieved from http://jalt-publications.org/proceedings/issues/2014-08_2013.1
- Baddeley, A. D. (1990). *Human memory: theory and practice*. Hove: Erlbaum.
- Cohen, A. D. (1993). *Language learning: insights for learners, teachers, and researchers*. Boston, MA: Heinle & Heinle.
- Coxhead, A. (2000). A New Academic Word List. *TESOL Quarterly*, 34(2), 213. doi: 10.2307/3587951
- Cross, D., & James, C. V. (2001). *A practical handbook of language teaching*. London: Longman.
- Elgort, I. (2010). Deliberate Learning and Vocabulary Acquisition in a Second Language. *Language Learning*, 61(2), 367-413. doi: 10.1111/j.1467-9922.2010.00613.x
- Gartner Your Source for Technology Research and Insight. (n.d.). Retrieved March 07, 2017, from <http://www.gartner.com/technology/home.jsp>.
- Hirschel, R., & Fritz, E. (2013). Learning vocabulary: CALL program versus vocabulary notebook. *System*, 41(3), 639-653.
- Horst, M., Cobb, T., & Meara, P. (1998). Beyond A Clockwork Orange: Acquiring second language vocabulary through reading. *Reading in a Foreign Language*, 11, 207-223.
- Hughes, A. (2013). *Testing for language teachers*. Cambridge: Cambridge University Press.
- Hulstijn, J. (2001). Intentional and incidental second language vocabulary learning: A reappraisal of elaboration, rehearsal, and automaticity. In P. J. Robinson (Ed.), *Cognition and second language instruction* (pp. 258-286). Cambridge: Cambridge University Press.
- Laufer, B., & Shmueli, K. (1997). Memorizing New Words: Does Teaching Have Anything To Do With It? *RELC Journal*, 28(1), 89-108. doi:10.1177/003368829702800106
- Lees, D. (2013). A Brief Comparison Of Digital- And Self-Made Word Cards For Vocabulary Learning. *Kwansei Gakuin University Humanities Review*, 18, 59-71. Retrieved June 2, 2017, from kwansai.repo.nii.ac.jp.
- Nakata, T. (2008). English vocabulary learning with word lists, word cards and computers: implications from cognitive psychology research for optimal spaced learning. *ReCALL*, 20(1), 3-20.
- Nation, I. S., & Webb, S. A. (2011). *Researching and analyzing vocabulary*. Boston, MA: Heinle, Cengage Learning.

- Nation, I. (2003). Effective ways of building vocabulary knowledge. *ESL Magazine*, 14-15.
- Nation, I. (2005). Language education: Vocabulary. In I. C. Brown (Ed.), *Encyclopaedia of language and linguistics* (2nd ed., Vol. 6, pp. 494-499). Oxford: Elsevier.
- Nation, I. (1995). Best practice in vocabulary teaching and learning. *EA Journal*, 7-15. Retrieved March 8, 2017, from <http://www.victoria.ac.nz/lals/about/staff/publications/paul-nation/1995-Best-practice.pdf>.
- Nikoopour, J., & Kazemi, A. (2014). Vocabulary Learning through Digitized & Non-digitized Flashcards Delivery. *Procedia - Social and Behavioral Sciences*, 98, 1366-1373.
- Puentedura, R. R. (2012, August 23). The SAMR Model: Background and Exemplars. Retrieved March 07, 2017, from http://www.hippasus.com/rrpweblog/archives/2012/08/23/SAMR_BackgroundExemplars.pdf.
- Quizlet. (2017). Retrieved March 07, 2017, from <https://quizlet.com>.
- Reinders, H., & White, C. (2011). The theory and practice of technology in materials development and task design. In N. Harwood (Ed.), *English language teaching materials: theory and practice* (pp. 58-80). Cambridge: Cambridge University.
- Richards, J. C., & Schmidt, R. W. (2002). *Dictionary of language teaching & applied linguistics*. Harlow: Longman.
- Soanes, C. (2010). *The paperback Oxford English dictionary*. Oxford: Oxford Univ. Press.
- Suppes, P., & Crothers, E. J. (1967). *Experiments in second-language learning*. New York: Academic Press.
- Webb, S. (2007). The Effects of Repetition on Vocabulary Knowledge. *Applied Linguistics*, 28(1), 46-65. doi:10.1093/applin/aml048
-

Research paper

Web 2.0 tools in the EFL classroom: Comparing the effects of Facebook and blogs on L2 writing and interaction

Gilbert Dizon* and Benjamin Thanyawatpokin**
Himeji Dokkyo University, Japan | Ritsumeikan University, Japan

*gdizon@gm.himeji-du.ac.jp | **btpokin@gmail.com

Abstract

Web 2.0 technologies have become an integral part of our lives, transforming not only how we communicate with others, but also how language is taught and learned in the L2 classroom. Several studies have looked into the use of these tools and how they influence L2 learning (e.g. Jin, 2015; Wang & Vásquez, 2014), yet only one has compared the effects of two Web 2.0 technologies (Castaneda Vise, 2008). Thus, the aim of this study was to compare the impact that Facebook and blogs had on the writing skills, namely, writing fluency, lexical richness, and syntactic complexity, of Japanese EFL learners. Moreover, the authors examined the influence blogging and Facebook had on interaction, i.e., the number of comments the learners posted outside of class. Student attitudes towards using these tools for written English were also measured through a survey based on the technology acceptance model (Davis, 1989). Twenty-three students at a Japanese university participated in the study and were divided into a Facebook group (n = 14) and a blog group (n = 9) according to their classes. Both groups took part in a ten-week treatment consisting of weekly guided free-writings on their respective Web 2.0 applications. Pre- and post-tests were administered and non-parametric statistical tests were used to determine if any significant writing gains were made. It was found that students in both blogging and Facebook groups showed similar improvements in writing skills. However, blogging seemed to be more effective at promoting interaction and students who took part in this group retained more favorable attitudes on using blogging for L2 writing. It was concluded that Facebook may indeed present an environment where students can be distracted from more formal educational pursuits (e.g. Wang & Kim, 2014) even when they are in private Facebook groups, while blogging may support a more serious environment for improving L2 writing skills.

Keywords: Web 2.0, computer-assisted language learning, L2 writing, EFL, Facebook, blogs.

1. Introduction

Education has entered an age in which there are almost countless numbers of ways in which students can use the Internet to practice or enhance their writing skills. Chief among them are Web 2.0 technologies, online tools which emphasize interaction, collaboration and creativity (Tu, Blocher, & Ntoruru, 2008). Compared with Web 1.0 tools such as email or traditional webpages, Web 2.0 applications are highly adaptable and promote greater degrees of interaction and collaboration (Harrison & Thomas, 2009; Pegrum, 2009). Upon the conception of these technologies, they were primarily used for recreational purposes (Crook, 2008), however through ingenuity in the classroom, they have become more and more relevant to the daily on-goings of language instruction. Researchers have attempted to analyze these tools and how applicable they are to L2 students. In their review of Web 2.0 literature, Wang and Vásquez (2012) found that the most cited advantage of these technologies is their ability to support a positive environment for language learning. Yet, they also note the lack of empirical research examining how Web 2.0 can impact language ability, specifically in regards to the less researched area of social-networking services (SNSs).

In recent years, researchers have tried to address this issue (e.g. Dizon, 2016; Shih, 2011; Wang & Vásquez, 2014), but studies incorporating comparison groups are still lacking, particularly those that compare the effects of different Web 2.0 tools on groups of learners. Students often engage in more than one type of Web 2.0 application; they may keep a Facebook profile for personal use but be asked to use more widely utilized forms of Web 2.0 such as blogs and wikis for formal language learning. Thus, it is imperative to investigate the degree to which each of these technologies affect student L2 proficiency. Therefore, this study fills this gap in the literature by comparing two technologies that represent the social and collaborative nature of Web 2.0 – blogs and Facebook – to see if there were any significant differences in writing improvements and interaction between two groups of Japanese EFL students using these tools, as well as assess the learners views towards their use in the L2 classroom.

2. Literature review

2.1. Facebook

Research shows that online interactions can promote communication among L2 learners (Kissau & Pyke, 2010; Moore & Iida, 2010), and this also seems to be true with Facebook. In a study involving L2 French students, Mills (2011) found that the SNS promoted social interaction, which in turn led to a sense of community among the participants. Jin (2015) had similar findings in a study investigating the effects of an intercultural exchange via Facebook between Korean EFL learners and American college students. Based on the quality as well as the high number of posts and comments, Jin (2015) concluded that Facebook fostered interaction and intercultural competence. In contrast to the previous studies, Alm (2015) examined the use of Facebook outside of formal language learning contexts. His findings indicated that the SNS had the potential to support engagement in the L2. However, what was key was whether or not a learner had native speaker friends on Facebook, illustrating the importance of native speakers as a language learning resource. Another interesting finding by Alm (2015) was that advanced language learners were more likely to use the SNS in their L2 as well as be part of a Facebook L2 group. These results illustrate that language students may use Facebook outside of the confines of the classroom as a means of authentic communication with others in the target language. Similarly, Mitchell (2010) found that the SNS encouraged the ESL college students in her study to communicate with their Facebook friends in English, thereby increasing their input as well as output in the L2.

L2 students seem to have mixed views towards the use of Facebook for language learning. Although all the learners in Alm's (2015) study stated that they felt less anxiety communicating in the L2 over Facebook, only advanced students viewed it as useful for informal language learning. Shih (2011) had similar findings in his study which focused on the use of Facebook and peer assessment with university EFL students in Taiwan. According to the results of questionnaires and interviews, Shih (2011) found that the participants had generally favorable opinions towards combining Facebook and peer assessment. In particular, the learners in his study thought the blended learning approach enhanced their L2 writing skills, reduced stress, and offered a convenient and fun way to communicate in the target language. However, a few downsides were listed by the participants as well, namely, the fact that writing through Facebook could lead to bad habits due to an over-reliance on online correction tools and the potential for the SNS to act as a distraction. The latter was also described as a disadvantage of Facebook in Wang and Kim's (2014) case study involving L2 Chinese students. Nonetheless, their overall perceptions towards Facebook were positive, and they indicated several benefits of using the site: low-pressure learning environment, opportunities to use Chinese, as well as strengthened relationships with their classmates. Likewise, the Malaysian EFL students in Kabilan, Ahmad, and Abidin's (2010) research stated a variety of advantages of Facebook writing including enhanced writing ability, confidence, motivation, and attitudes towards the L2. Kabilan et al. (2010) also had a few negative findings regarding Facebook. To be specific, some of the learners in the study thought that the SNS was not a suitable environment for studying English. Moreover, a few of them indicated that they could not improve their English skills through Facebook because it is merely a social space to share stories and information with friends.

While research indicates that Facebook can support improvements in L2 writing output or fluency, it is unclear if the SNS can enhance the quality of students' writing. Wang and Vásquez (2014) examined the use of Facebook with L2 Chinese learners and found that the Facebook group in their study wrote significantly more Chinese characters on a post-test than a control group which took part in no treatment. In regards to writing quality however, there were no significant differences between the two groups. Similar results were found by (Dizon, 2016) in a study of Facebook and EFL learners. Two groups were involved in the study, an experimental group which used Facebook and a comparison group which used paper-and-pencil writing. Although the Facebook group improved their writing fluency to a greater degree than the comparison group on a timed post-test, significant differences in lexical richness or grammatical accuracy were not found. To date, only Shih's (2011) study has found that the use of Facebook could lead to significant improvements in L2 writing quality. However, the study incorporated a combination of peer assessment and Facebook and no control group was involved. Therefore, it is unknown if Facebook truly had a significant impact on the improvements that were made or if peer assessment was a greater factor.

2.2. Blogging

Blogging and its uses in the classroom is one of the most used, and widely researched, applications to arise out of the Web 2.0 era (Wang & Vásquez, 2012). Many studies have made note of the benefits involved with using blogs in a language learning classroom (e.g. Sykes, Oskoz & Thorne, 2008; Warschauer, 2010). Similar to Facebook, blogging has also been shown to promote communication amongst students; this is reported to be achieved by the propensity of blogging to facilitate communication amongst students using it to write. In a study done by Nepomuceno (2011), ESL students enrolled in academic writing classes were observed to comment on blog posts made by their classmates on a number of different topics. These topics were not limited by the researcher and students took it upon their own volition to reach out to their classmates and communicate. In addition, students stated that through blogging they were able to make new friends. Pinkman (2005) also found similar results in an action-research study she conducted using blogs with Japanese EFL students. Commenting on blogs prompted students to spend time thinking up ways to respond to their peers. People who were not involved with the class also began commenting on the student blogs which led to the conclusion that blogs being a public open forum could potentially invite more global communication and encourage students to speak with people in other countries through this medium. Similar to the Alm's (2015) findings on Facebook, Hashimoto (2012) found that proficient L2 students were able to experience authentic communication in their target language through the Web 2.0 tool, thereby promoting learning autonomy.

As a whole, blogs appear to be welcomed by students when used for written exercises or activities; however, the extent to which students enjoy blogging seems to vary by student and certain environmental factors (e.g. Miyazoe & Anderson, 2009; Pinkman, 2005). Nepomuceno (2011) stated that his students felt blogging was an extremely positive experience in that it allowed them to be connected to the wide world in addition to being convenient when compared to other methods of writing for English classes. In a study done in the Japanese EFL classroom, Miyazoe and Anderson (2009) observed that students who used blogging in class in conjunction with other Web 2.0 applications such as Wikis and BBS had a positive impression of it in class. The overall conception of blogs may have been positive, but the actual usage of the blogs in the classroom demonstrated a noted lack of commenting on each other's blogs. In fact, the researchers stated that the students felt that each blog was a "private space" in which they were allowed to post their thoughts. It should be noted, however, that students did in fact read their classmates' blogs. In another study, Amir, Ismail, and Hussin (2011) collected data which suggested students became more motivated after using blogs to write about certain topics for a period of six weeks. Survey data revealed students felt an increased interest in writing which helped improve their writing skills and also they had more confidence in their writing after the study was concluded.

In terms of language-related outcomes, research has shown that students have gained certain benefits from using blogs in the classroom. Fellner and Apple (2006) observed

that low-level Japanese students writing blogs in a seven-day 20-minute blog activity achieved higher writing fluency scores. In addition to writing faster in the allotted time, a lexical analysis of the students' blogs revealed that more complex words were also being used which meant that lexical complexity was also promoted. Nakatsukasa (2009) found similar results when comparing the lexical complexity of students' blog posts after a set amount of timed blogging sessions in an ESL classroom. Students were observed to use more complex words after several weeks of completing an assignment where they had to blog collaboratively and comment on each other's posts. However, post length was found to be determined by student interest in the topic and not simply by becoming accustomed to the process of blog writing or incidental learning. In a study done with Japanese students using Moodle, Miyazoe and Anderson (2009) observed that students had higher levels of lexical density, or ratio of total words to different words used in the text. The researchers also claimed that there were higher levels of complexity and vocabulary usage. However, these claims were not subjected to any quantitative analysis. To date, there have not been many studies that quantitatively attempt to rate the syntactic complexity of student texts after writing in blogs for an extended period of time.

In summary, Facebook and blogs have strikingly similar benefits and have been implemented into more and more curricula in recent years. Both have been shown to promote interaction among L2 students (Alm; 2015; Jin, 2015; Mills, 2011; Mitchell, 2010; Nepomuceno, 2011; Pinkman, 2005) as well as authentic L2 communication (Alm, 2015; Hashimoto, 2012). In addition, learners seem to have generally positive views towards their use for L2 learning, despite the fact that these web 2.0 tools come with their own unique disadvantages as well (Alm, 2015; Amir et al., 2011; Kabilan et al., 2010; Miyazoe & Anderson, 2009; Nepomuceno, 2011; Shih, 2011; Wang & Kim, 2014). Lastly, research on Facebook and blogs have indicated that students can make L2 language gains through their use, particularly when it comes to writing output and vocabulary improvements (Dizon, 2016; Fellner & Apple, 2006; Miyazoe & Anderson, 2009; Nakatsukasa, 2009; Shih, 2011; Wang & Vásquez, 2014). However, while studies like the ones listed above have delved into using different forms of computer-mediated communication (CMC), few have compared two different web 2.0 technologies. The lone exception to this is Castaneda Vise's (2008) study which compared two groups of L2 Spanish learners: one which used wikis and another which used blogs. The researcher found that there were no significant differences between the groups in terms of achievement and satisfaction levels. Despite these findings, much more research needs to be done in order to determine which Web 2.0 tool is more suitable in formal language learning contexts. Therefore, this study aims to fill this gap in the literature by addressing the following research questions:

- Are there significant differences in writing fluency, lexical richness, or syntactic complexity between students who use Facebook or blogs?
- Is there a significant difference in student interaction between the two groups?
- Are there differences in students' opinions of each web 2.0 tool?

3. Methodology

3.1. Research design

A mixed method quasi-experimental design was implemented in this study to examine if there were any significant differences in writing output, lexical richness, or syntactic complexity, between two groups of students: one which wrote in class via Facebook and another which used blogs. Two writing assessments were administered at the start and completion of the treatment to measure if any language improvements were made in these three areas. In addition, the number of student comments in the Facebook and blog groups were recorded to see if there was a significant difference in the amount of interaction between the members of each group. The qualitative aim of the study was to survey each group's attitudes towards Facebook and blogs with the technology acceptance model (Davis, 1989) in order to determine if there were any differences between their opinions in regards to their perceived usefulness, perceived ease of use, and behavioral intention to use the Web 2.0 tools.

3.2. Participants

A total of 23 first- and second-year EFL students at a small, private Japanese university agreed to participate in the study. The learners were part of the Department of Humanities and Social Sciences at the university and were enrolled in a course entitled Communicative English. The students were divided into four separate classes based on their English test scores in a Japanese standardized test called *Eiken*. The classes met three times a week in 90-minute periods during the spring 2017 semester. As shown in Table 1 below, each researcher taught a Facebook and a blog group as well as one first-year and second-year class in order to minimize the impact of teacher effects on the results of the study.

Group	Researcher		Year	
	A	B	1st	2nd
Facebook (n = 14)	8	6	6	8
Blog (n = 9)	5	4	5	4

3.3. Treatment

While the students using Facebook posted their writing on separate group pages, those in the blog group wrote on individual class pages created with Blogger.com since the students already had Gmail accounts through the university. Both groups took part in a 10-week treatment consisting of 15-minute guided freewritings (GFs). Whereas traditional freewriting involves students writing about a topic of their choice, GF is more focused, which encourages learners to get the writing process started, a common difficulty among EFL writers (Hwang, 2010). Additionally as Hammond (1991) asserts, GF better develops students' critical thinking skills, especially when writing is shared; thus making Web 2.0 technologies a natural medium for GF. Writing topics were selected by the researchers and were the same for the Facebook and blog groups in order to maintain consistency in writing themes between the classes involved (Table 2). At the start of each GF, topics were introduced and a short writing prompt was provided to the students. Other than this, no additional guidance was provided for the students; they were not allowed to use their electronic dictionaries, smartphones, or any other writing aid. After completing each GF, the students in both groups were assigned to comment on at least two other posts on their respective Facebook group pages or blogs outside of class in order to promote interaction among the learners. It was also highly encouraged to comment more than twice. It is also important to note that all the Facebook and blog pages were set to the strictest privacy settings. This was done to maintain the privacy of the students and to prevent outside interaction, which could possibly damage student motivation as was witnessed in Pinkman (2005).

Week	Topic
Week 1	Travel
Week 2	Work
Week 3	Health
Week 4	University life
Week 5	Foreign language

Week 6	Environment
Week 7	Hobbies
Week 8	Money
Week 9	Education
Week 10	Talents

3.4. Research instruments

While the students wrote about several different topics during the treatment period, pre-and post-tests were used in order to assess if the students' writing improved. The writing procedure for the assessments was identical to the treatment. The pre- and post-test both used topics that asked the students to report on their plans for certain school holidays. The pre-test focused on the students' plans for Golden Week (a five-day holiday in Japan) and the post-test asked the students about their plans for the summer vacation. The themes of the assessments were kept similar due to the effect that topic can have on lexical richness (Laufer & Nation, 1995; Robinson, 2001).

A 10-item, L1 questionnaire based on the technology acceptance model (TAM), developed by Davis (1989), was created by the researchers and administered to assess each group's views of Facebook and blogs. According to Lee, Kozar, and Larsen (2003), TAM is "the most influential and commonly employed theory for describing an individual's acceptance of information systems" (p. 752). Although TAM can also measure external factors such as user training and anxiety, it consists of three primary variables: perceived usefulness (PU), perceived ease of use (PEOU) and behavioral intention (BI) (Figure 1). Davis (1989) defined PU as the degree of usefulness that a particular system or technology provides which enables a person to improve their performance, while PEOU is the degree to which a person believes a given technology can be used without effort. Eight of the survey items, four each, pertained to PU and PEOU, with the remaining two related to BI. The students were asked to rate their level of agreement towards the questionnaire items according to a 5-point Likert scale. The reliability of the survey was verified with Cronbach's alpha (α), with sub-scale values all > 0.8 (PU = .844; PEOU = .865; BI = .814), indicating a good level of internal consistency (George & Mallery, 2003).

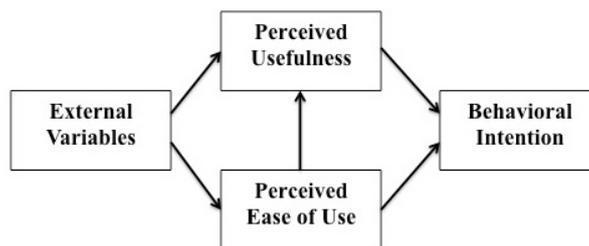


Figure 1. The technology acceptance model.

3.5. Variables

The independent variables in this study were the two web 2.0 tools that were investigated: Facebook and blogs. A total of three dependent variables were examined to measure any writing gains that were made: (1) writing fluency, or the number of words written on the pre- and post-tests, (2) lexical richness, i.e., the ratio of words

written beyond level 1 of the New General Service List (Browne, 2013) and the words written on the writing assessments, and (3) syntactic complexity, measured by subordination as was demonstrated by Nation (1989). For writing fluency, it was decided to measure word count instead of syllable count in order to keep in line with conventions established by other researchers (e.g. Nakatsukasa, 2009). The average number of comments made per week by each student outside of class was also examined to see if there was a significant difference in the level of interaction between the learners of each group. Qualitatively, the students' attitudes towards the PU, PEOU, and BI to use blogs or Facebook for English writing were measured to assess their views of the Web 2.0 technologies.

3.6. Data collection and analysis

The bulk of the data for the study was collected from the Facebook and Blogger class pages. The comments that the students posted on these two websites were used to gather data to determine the interaction level of the learners. Both pre- and post-tests were conducted on the respective websites in order to keep in the theme of facilitating a class-based activity centered on CMC. In order to gauge student perceptions of Facebook or Blogger, the students also completed a TAM-based online survey via surveymonkey.com; the contents of which asked students various questions about the PU, PEOU, and BI of the Web 2.0 tools based on a 5-point Likert scale.

Lexical richness data was analyzed using the New General Service List (NGSL) version of VocabProfile, an online vocabulary profiler based on Laufer and Nation's (1995) Lexical Frequency Profile. This version of VocabProfile breaks down English text into five categories: the first three levels of the NGSL, which is comprised of the 2,801 most important highly frequently used words in the English language (Browne, 2013), the New Academic Word List, and off-list vocabulary, words that do not fall into any of the previously mentioned categories. Because of this, proper nouns as well as non-English words were removed prior to analysis as their inclusion would have skewed the results of the analysis.

Non-parametric statistical tests were employed to analyze the pre-and post-test writing data due to the small sample size. The Sign test was used to determine if significant writing improvements were made within each group while the Mann-Whitney U test was used to assess if there were significant differences between the groups in relation to any gains which were made. The latter was also used to determine if there was significant difference in the interaction level, i.e., the average number of student comments, between the learners who used Facebook and those that wrote on blogs. Descriptive statistics of the survey data, detailing the mean and SD values of the each survey item as well as the TAM constructs, were provided to illustrate the learners' views of Facebook or blogs.

4. Results

4.1. Writing fluency

Results from the Mann-Whitney U test indicate that the writing fluency of the Facebook group (Mdn = 45) did not significantly differ from the blog group (Mdn = 60) on the pre-test, $U = 47.5$, $p = .483$. In other words, the groups were equivalent in terms of writing fluency ability prior to the start of the treatment. Both the Facebook ($Z = 3.05$, $p = .002$) and the blog group ($Z = 2.33$, $p = .019$) showed significant increases in average word count from pre- to post-test, with the blog group making larger gains. However, the improvements between the Facebook group (Mdn = 19) and the blog group (Mdn = 30) were not significantly different, $U = 48.5$, $p = .528$.

Table 3. Writing fluency results				
Group	Pre-test		Post-test	
	Mean WC	SD	Mean WC	SD
Facebook	53.15	30.07	74.38	43.90
Blog	60.78	19.41	96.78	41.03

NB: WC = word count

4.2. Lexical richness

Similar to writing fluency, a significant difference was not found between the Facebook group (Mdn = .071) and the blog group (.054) as it pertains to lexical richness at the outset of the study, $U = 42.5$, $p = .298$. As shown below in Table 3, both groups were able to make slight gains from the pre-test to the post-test. However, the improvements made within the Facebook group ($Z = 1.94$, $p = .052$) and the blog group ($Z = 1.66$, $p = .095$) were not significant. Moreover, the difference between each group's gains (Facebook Mdn = .026, Blog Mdn = .034) were not found to be significant, $U = 57$, $p = .944$.

Table 4. Lexical richness results				
Group	Pre-test		Post-test	
	Mean LR	SD	Mean LR	SD
Facebook	6.5%	3.7%	9.4%	4.0%
Blog	5.2%	2.7%	8.4%	3.5%

NB: LR = lexical richness

4.3. Syntactic complexity

According to the Mann-Whitney U test, the Facebook (Mdn = 1) and blogging (Mdn = 3) groups were statistically different for data taken during the pre-test ($U = 24.5$, $p = .025$). Essentially, this means that the blogging group was using more subordinate clauses than the Facebook group from the outset of the study. The Sign test showed the Facebook group ($Z = 2.53$, $p = .011$) exhibited significant improvements in their use of clauses. In addition, the blogging group ($Z = 2.33$, $p = .020$) also showed scores that signify significant improvement in syntactic complexity. When data from the Facebook (Mdn = 1) and blog group (Mdn = 3) are compared, though, there does not seem to be a significant difference between them, $U = 37$, $p = .162$.

Table 5. Syntactic complexity results				
Group	Pre-test		Post-test	
	Mean SC	SD	Mean SC	SD
Facebook	3.11	1.27	6.22	1.72
Blog	1.62	1.33	3.23	2.13

NB: SR = syntactic complexity

4.4. Interaction

As explained above, the interaction between students was measured by the average number of times they commented on posts every week. The Facebook group (Mdn = 1.69) and the blog group (Mdn = 2.17) were found to be statistically different, $U = 22$, $p = .038$. In other words, the blogging group was observed to be creating significantly more comments on average each week for the duration of the study.

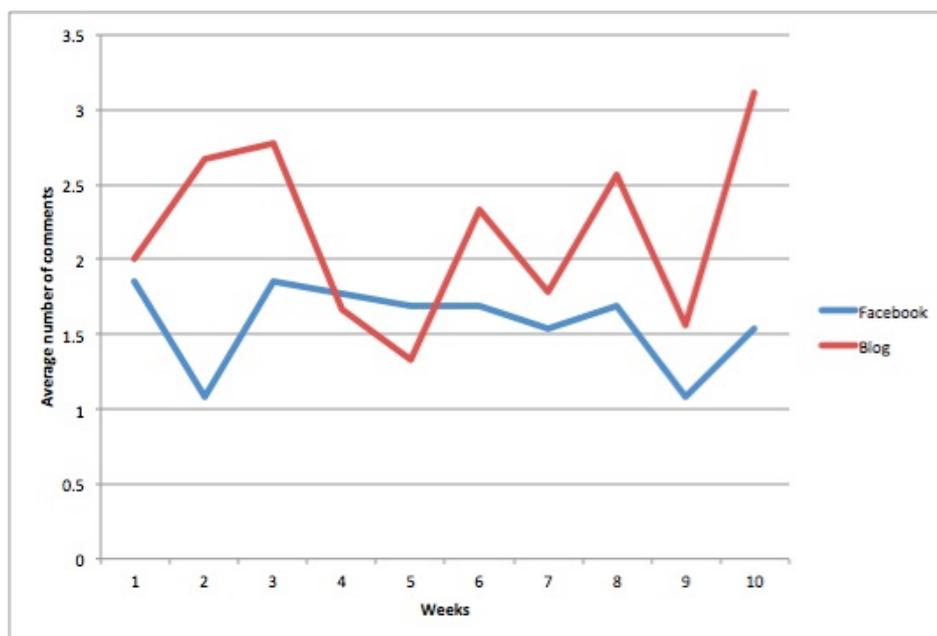


Figure 2. Average number of comments per week.

From the line graph above, it can be observed that both groups started out week one of the study with comment counts that were quite similar. However, during the first half of the treatment, the number of comments from the blog group fluctuated, rising or falling behind the Facebook group sporadically. Over the last five weeks of the study, the students in the blog group consistently commented more often than those in the Facebook group. Interestingly, the blog group showed more fluctuations in the number of comments posted, but the Facebook group comments stayed relatively even.

4.5. Student attitudes

As shown in Table 5, the blog group had higher levels of agreement on eight out of the ten items of the questionnaire. Six of these items had an agreement rating of 4.0 or higher, whereas the Facebook group rated one statement (PU2) with a similar level of agreement. Accordingly, the blog group also had higher levels of agreement on each survey construct: PU, PEOU, and BI. The only item which rated higher for the Facebook

group was statement one, which related to writing speed. The first item related to PEOU resulted in the lone tie, with both groups having the same level of agreement towards the ease of use to write on Facebook/blogs.

#	Statement	Facebook		Blog	
		Mean	SD	Mean	SD
PU1	I was able to write more quickly on Facebook/the blog.	3.92	0.95	3.56	0.53
PU2	Writing on Facebook/the blog improved my writing performance.	4.00	0.71	4.33	0.71
PU3	Writing on Facebook/the blog made it easier to write in English.	3.77	0.93	4.22	0.67
PU4	Facebook/Blog writing was useful in my class.	3.92	0.76	4.00	0.71
PU Mean		3.90	0.82	4.03	0.82
PEOU1	It was easy for me to write on Facebook/the blog.	3.00	1.08	3.00	0.50
PEOU2	It was easy for me to become skillful at writing on Facebook/the blog.	2.92	1.04	3.33	0.71
PEOU3	Learning how to write on Facebook/the blog was easy for me.	3.54	1.05	4.00	0.50
PEOU4	The Facebook group/class blog page was clear and understandable.	3.69	0.95	4.11	0.60
PEOU mean		3.29	1.05	3.61	1.05
BI1	I intend to take more classes using Facebook/blog writing in the future.	3.54	1.27	4.22	0.44
BI2	If I am offered, I intend to write more English posts and comments on Facebook/blogs.	3.31	1.03	3.78	0.67
BI mean		3.42	1.14	4.00	0.59

5. Discussion

5.1. Are there significant differences in writing fluency, lexical richness, or syntactic complexity between students who use Facebook or blogs?

All groups who took part in the study were observed to have statistically significant gains in writing fluency and syntactic complexity. These findings are in line with previous literature on Web 2.0 tools in terms of writing output (Dizon, 2016; Fellner & Apple, 2006; Wang & Vásquez, 2014). However, when the level of improvements were compared, there were no significant differences. In regards to lexical richness, both

groups did not make significant improvements, which contradicts the lexical gains made in other studies (Fellner & Apple, 2006; Miyazoe & Anderson, 2009; Nakatsukasa, 2009). Judging by this data, it can be concluded that both writing on Facebook and blogging had the same level of effect on the students. In other words, the methods were equally as effective when analyzing the written texts that students produced in an EFL setting.

5.2. Is there a significant difference in student interaction between the two groups?

Although interaction was promoted through the use of Web 2.0 technology, as indicated by the average number of comments made per week in both groups, blogging spurred students to interact with their fellow classmates more significantly than Facebook. This confirms the positive role that blogging has on communication between L2 learners (Nepomuceno, 2011; Pinkman, 2005), and casts some doubt on the effect of Facebook, as other researchers have lauded the impact of the SNS on student interaction (Alm, 2015; Jin, 2015; Mills, 2011; Mitchell, 2010). While Facebook may come with several benefits, it also comes with its own set of downsides for L2 students, including the fact that it has the potential to distract students from learning outcomes (Shih, 2011; Wang & Kim, 2014), which could have affected the learners in the Facebook group when commenting outside of class.

5.3. Are there differences in students' opinions of each web 2.0 tool?

Students' perceptions towards the use of Facebook and blogs were generally positive, reinforcing past findings on Web 2.0 tools (e.g., Nepomuceno, 2011; Shih, 2011). However, the blog group had higher levels of agreement towards PU, PEOU, and BI, which suggests that L2 students may prefer blogging over Facebook writing. As found by Kabilan et al. (2010), the overt social and recreational nature of Facebook may turn off some students to L2 writing on the SNS. In contrast, the blogs in the study were created and used for the sole purpose of improving English writing. Therefore, language instructors ought to train students on how to best leverage the features of Facebook for language learning purposes in order for it to be used effectively.

6. Limitations

First and foremost, one of the study's obvious limitations is the small sample size. Studies done by other researchers have used around 20 to 30 students for one class (e.g. Fellner and Apple, 2006; Nepomuceno, 2011). If this number were to be used as a baseline, then it could be considered that the current study did not gather enough participants to provide results that can be generalized to a larger student population. Non-parametric statistical analysis of the data was also used, which was not utilized by most previous studies. This could have also affected the generalizability of the data.

A further limitation was the fact that the two groups started the study at different levels of English as it pertains to syntactic complexity. In response to this, we decided to focus on the extent to which each group was improving throughout the course of the study. Other studies, such as Fellner and Apple (2006) used students that had a variety of English comprehension levels. Montero-Fleta and Perez-Sabater (2010) also included participants that were described to include students of English language comprehension levels.

While we included the number of blog and Facebook comments in the data analysis, we did not include the content of these comments. Other studies such as Nakatsukasa (2009) went in-depth into the linguistic makeup of these comments that could also help suggest English improvements. In further studies, it would be beneficial to look into what students are posting over the course of several weeks.

Lastly, some researchers advocate that broad comparisons, such as the one performed in this study, ought to be avoided (Levy & Stockwell, 2006). Instead, Levy and Stockwell (2006) recommend that more narrowly defined comparative studies be done in order to focus on the specific design features that help facilitate effective language learning. Given this, it would be worthwhile to examine whether design features within Web 2.0 tools such as privacy settings or automatic translation have significant effects on promoting L2 skills and interaction.

7. Conclusion

The primary goal of this study was to compare the influence that Facebook and blogs can have on L2 writing and interaction. In terms of writing improvements, 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. These findings complement previous research and further demonstrate that Facebook and blogs both have the potential to improve the writing skills of L2 learners (Dizon, 2016; Fellner & Apple, 2006; Miyazoe & Anderson, 2009; Nakatsukasa, 2009; Shih, 2011; Wang & Vásquez, 2014). Evidence suggested that blogs were superior at enhancing interaction between the learners, indicating that the social and recreational nature of Facebook may distract learners from their language goals (Shih, 2011; Wang & Kim, 2014). Lastly, the data from the TAM-based questionnaire signified that the students in the blog group had more positive views towards the use of the Web 2.0 technology than those in the Facebook group. Although more empirical research on Facebook and L2 learning has been done in recent years, the use of blogs is still much more established in the field (Wang & Vásquez, 2012). Consequently, some students may be wary of the use of Facebook in formal learning contexts, especially since it is primarily used for recreation and non-academic communication. Given this, language instructors must carefully examine their own teaching contexts, placing special emphasis on the needs, abilities, and resources of their learners, before deciding to implement any Web 2.0 tool in the classroom.

References

- Alm, A. (2015). Facebook for informal language learning. *EUROCALL Review*, 23(2), 3-18. Retrieved from <http://polipapers.upv.es/index.php/eurocall/article/view/4665/4791>
- Amir, Z., Ismail, K., Hussin, S. (2011). Blogs in language Learning: Maximizing Students' Collaborative Writing. *Procedia Social and Behavioral Sciences*, 18, 537-543. doi: 10.1016/j.sbspro.2011.05.079
- Browne, C. (2013). The new general service list: Celebrating 60 years of vocabulary learning. *The Language Teacher*, 37(4), 13-16.
- Castaneda Vise, D. A. (2008). The effects of wiki- and blog-technologies on the students' performance when learning the preterit and imperfect aspects in Spanish. *Dissertation Abstracts International: Section A. The Humanities and Social Sciences*, 69 (01), 0187.
- Crook, C. (2008). *Web 2.0 Technologies for Learning: The Current Landscape – Opportunities, Challenges and Tensions*. Retrieved from http://dera.ioe.ac.uk/1474/1/becta_2008_web2_currentlandscape_litrev.pdf
- Davis, F. D. (1989). Perceived usefulness, perceived ease of use, and user acceptance of information technology. *MIS Quarterly*, 13(3), 319-340. Retrieved from jstor.org/stable/249008
- Dizon, G. (2016). A comparative study of Facebook vs. paper-and-pencil writing to improve L2 writing skills. *Computer Assisted Language Learning*, 29(8), 1249-1258. doi: 10.1080/09588221.2016.1266369
- Fellner, T., Apple, M. (2006). Developing writing fluency and lexical complexity with blogs. *The JALT CALL Journal*, 2(1), 15-26.
- George, D., & Mallery, P. (2003). *SPSS for Windows step by step: A simple guide and reference* (4th ed.). Boston, MA: Allyn & Bacon.
- Hammond, L. (1991). Using focused freewriting to promote critical thinking. In P. Belanoff, P. Elbow, & S. Fontaine (Eds.), *Nothing begins with N: New investigations of freewriting* (pp. 71-92). Carbondale, IL: Southern Illinois University.
- Harrison, R., & Thomas, M. (2009). Identity in online communities: social networking sites and language learning. *International Journal of Emerging Technologies and Society*, 7(2), 109-124.

- Hashimoto, K. (2012). *Exploring the relationship between L2 blogging, learner autonomy, and L2 proficiency levels: A case study of post-secondary Japanese L2 learners* (Doctoral dissertation). Retrieved from Proquest, UMI Dissertations Publishing.
- Hwang, J. A. (2010). A case study of the influence of freewriting on writing fluency and confidence of EFL college-level students. *Second Language Studies*, 28(2), 97–134. Retrieved from <http://www.hawaii.edu/sls/wp-content/uploads/2014/09/Hwang.pdf>
- Jin, S. (2015). Using Facebook to promote Korean EFL learners' intercultural competence. *Language, Learning & Technology*, 19(3), 38–51. Retrieved from <http://llt.msu.edu/issues/october2015/action2.pdf>
- Kabilan, M. K., Ahmad, N., & Abidin, M. J. Z. (2010). Facebook: An online environment for learning of English in institutions of higher education? *Internet and Higher Education*, 13, 179-187. doi: 10.1016/j.iheduc.2010.07.003
- Kissau, S., McCullough, H., Pyke, J. G. (2010). 'Leveling the playing field:' The effects of online second language instruction on student willingness to communicate in French. *CALICO Journal*, 27(2). 277-297. doi: 10.11139/cj.27.2.277-297
- Laufer, B., & Nation, P. (1995). Vocabulary size and use: Lexical richness in L2 written production. *Applied Linguistics*, 16, 307–322. doi:10.1093/applin/16.3.307
- Lee, Y., Kozar, K. A., & Larsen, K. R. T. (2003). The technology acceptance model: Past, present, and future. *Communications of the Association for Information Systems*, 12, 752-780. Retrieved from aisel.aisnet.org/cais/vol12/iss1/50
- Levy, M., & Stockwell, G. (2006). *CALL Dimensions: Options and issues in CALL*. Lawrence Erlbaum: Mahwah, NJ.
- Mills, N. (2011). Situated learning through social networking communities: The development of joint enterprise, mutual engagement, and a shared repertoire. *CALICO Journal*, 28(2), 345-368. doi: 10.11139/cj.28.2.345-368
- Mitchell, K. (2010). A social tool: Why and how ESOL students use Facebook. *CALICO Journal*, 29(3), 471-493. doi: 10.11139/cj.29.3.471-493.
- Miyazoe, T., Anderson, T. (2009). Learning outcomes and students' perceptions of online writing: Simultaneous implementation of a forum, blog, and wiki in an EFL blended learning setting. *System*, 38, 185-199.
- Montero-Fleta, B., & Perez-Sabater, C. (2010). A research on blogging as a platform to enhance language skills. *Procedia Social and Behavioral Sciences*, 2, 773-777. doi: 10.1016/j.sbspro.2010.03.100
- Moore, K. & Iida, S. (2010). Students' perception of supplementary, online activities for Japanese language learning: Groupwork, quiz and discussion tools. *Australasian Journal of Educational Technology*, 26(7), 966-979. Retrieved from <http://www.ascilite.org.au/ajet/ajet26/moore.html>
- Nakatsukasa, K. (2009). The efficacy and students' perceptions of collaborative blogging in an ESL classroom. In Chapelle, C.A., Jun, H.G., Katz, I (Eds.), *Developing and evaluating language learning materials* (pp. 69-84). Ames, IA: Iowa State University.
- Nation, I. S. P. (1989). Improving speaking fluency. *System*, 17, 377–384. doi: 10.1016/0346-251X(89)90010-9
- Nepomuceno, M. (2011). Writing Online: Using Blogs as an Alternative Writing Activity in Tertiary ESL Classes. *TESOL Journal*, 5, 92-105.
- Pegrum, M. (2009). Communicative networking and linguistic mashups on Web 2.0. In M. Thomas (Ed.), *Handbook of research on Web 2.0 and second language learning* (pp. 20-41). IGI Global.
- Pinkman, K. (2005). Using Blogs in the Foreign Language Classroom. *The JALT CALL Journal*. 1(1), 12-24.
- Robinson, P. (2001). Task complexity, task difficulty, and task production: Exploring interaction in a componential framework. *Applied Linguistics*, 22(1), 27–57. doi: 10.1093/applin/22.1.27

Shih, R. C. (2011). Can Web 2.0 technology assist college students in learning English writing? Integrating Facebook and peer assessment with blended learning. *Australasian Journal of Educational Technology*, 27(5), 829-845. Retrieved from <https://eric.ed.gov/?id=EJ943540>

Sykes, J. M., Oskoz, A., & Thorne, S. L. (2008). Web 2.0, Synthetic Immersive Environments, and Mobile Resources for Language Education. *CALICO Journal*, 25(3), 528-546.

Tu, C., Blocher, M., & Ntoruru, G. (2008). Constructs for Web 2.0 learning environments: A theatrical metaphor. *Educational Media International*, 45(4), 253–269. doi: 10.1080/09523980802588576

Wang, S., & Kim, D. (2014). Incorporating Facebook in an intermediate-level Chinese language course: A case study. *IALLT Journal*, 44(1), 38-78. Retrieved from http://old.iallt.org/iallt_journal/incorporating_facebook_in_an_intermediate_level_chinese_language_course_a_case_study

Wang, S., & Vásquez, C. (2012). Web 2.0 and second language learning: What does the research tell us? *CALICO Journal*, 29(3), 412–430. doi: 10.11139/cj.29.3.412-430

Wang, S., & Vásquez, C. (2014). The effect of target language use in social media on intermediate-level Chinese language learners' writing performance. *CALICO Journal*, 31(1), p. 78-102. doi: 10.11139/cj.28.2.345-368

Warschauer, M. (2010). Invited Commentary: New Tools for Teaching Writing. *Language Learning and Technology*, 14(1), 3-8. Retrieved from <http://llt.msu.edu/vol14num1/commentary.pdf>

Reflective practice

Digital video creation in the LSP classroom

Ornaith Rodgers* and Labhaoise Ni Dhonnchadha**
National University of Ireland, Galway, Ireland

*ornaith.nidhuibhir@nuigalway.ie | **labhaoise.nidhonnchadha@nuigalway.ie

Abstract

The twenty-first century world of digital media and multimodalities demands a rethinking of approaches to languages for specific purposes (LSP). This article seeks to determine the effectiveness of digital video creation as a teaching and learning tool in the LSP context through an investigation of students' perceptions of the usefulness of this activity. The study is based on a digital video creation project carried out with a group of second year undergraduate students on the BSc in Biotechnology programme in NUI Galway who also study French as part of their degree programme. The findings are indicative of an overwhelmingly positive response from learners to this activity, both in terms of the development of language skills and other key social and professional skills. However, findings also warn that students' digital competencies must not be over-estimated, despite a general assumption in technology-enhanced language learning research, that the current generation of students have a high level of digital literacy. This study highlights the pedagogical potential of digital video creation in the language classroom and demonstrates that it embraces many of the core elements underpinning progressive LSP pedagogy, by giving students the opportunity to keep pace with the multimodality afforded by digital media and by ensuring their language learning is both contextualised and authentic. It advocates the use of digital video creation in language learning and particularly in LSP, by highlighting the strong impact that this activity had on the participants in this study.

Keywords: Languages for Specific Purposes, video, language pedagogy, digital literacy, multimodality.

1. Introduction

Since the 1960s, advances in instructional technologies have changed the landscape of foreign language teaching and learning by providing new possibilities for learning in ways beyond sitting in a traditional classroom (Duman, Orhon & Gedik, 2014, p. 197). Most language classes are taught using the support of computer-based multimedia in the form of audio, graphics or video and the internet is also regularly used for language learning (Burston, 2016: 3). However, during the last 10 years, the widespread ownership of mobile technologies such as smartphones, media players and tablet computers has encouraged a new dimension to technology-enhanced learning. While Mobile-Assisted-Language Learning (MALL) has been in existence for over 20 years, improvements in connectivity, Bluetooth, GPRS, storage and processing have extended the capabilities of mobile devices to tools that can be used to facilitate language learning (Duman et al, 2014: 198). Developments in computer assisted language learning (CALL), MALL and computer-mediated communication (CMC) have combined to transform the language classroom into what can be termed "a fertile venue for testing out innovative technology-based projects aimed at empowering language learners" (Dugartsyrenova & Sardegna, 2016: 59). It is against this backdrop of advances in instructional technology, that the role of video in the language classroom needs to be re-evaluated.

Advances in digital technology have created exciting opportunities not just for language learning in general, but particularly for dynamic uses of video in the language classroom. Students are living in a society where the use of technology is an integral

aspect of everyday living and are literate in ways that differ from previous generations. Prensky (2001) distinguished between *digital natives* (born into the digital era) and *digital immigrants* (those who grew up in the pre-digital era). Students are arguably *digital natives* (Prensky & Heppell, 2008), capable of dealing with multi-modal and digital texts which require non-sequential processing (Dal, 2010: 2). Although already an integral part of foreign language teaching, digital technology is destined to play an increasing role in language teaching in the coming years. As we move further into the 21st century, the distinction between *digital natives* and *digital immigrants* becomes less relevant, and Prensky (2012, 181) emphasises the need for *digital wisdom*, arguing that digital technology can make us wiser, that “it is from the interaction of the human mind and digital technology that the digitally wise person is coming to be” (Prensky, 2012: 182). He further insists that “educators are digitally wise when they let students learn by using new technologies, putting themselves in the roles of guides, context providers, and quality controllers” (Prensky, 2012: 190). It is also within the context of the need for *digital wisdom* that this study arises.

This study focusses on the practice of digital video creation in the language classroom with a specific focus on the Languages for Specific Purposes (LSP) classroom. Naqvi and Mahrooqi (2016: 49) use the abbreviation of SCDV (student created digital video) for this practice and define it as follows:

SCDV refers to the practice where students, either individually or in groups take part in the creation of a short video using either online software programs or their own software and hardware”. Students engage in researching, recording, directing, storyboarding, scripting, practicing and performing (...), editing and other post-production activities.

In this study, we examine the effectiveness of digital video creation as a teaching and learning activity in the context of LSP. This is done through an analysis of students’ perceptions of the use of digital video creation as a tool to enhance language learning in the LSP classroom, and a comparison of these initial findings with an analysis of the videos created. The article first reviews literature in the related fields of video in the language classroom and LSP, before outlining the factors which point to the pedagogical potential of digital video creation in the LSP field. The study itself is based on a digital video creation project carried out with a group of second year University students in Galway, Ireland who are studying French as part of their degree in Biotechnology, and is outlined in terms of context, participants, methodology and analysis of results, focusing on important findings and their relevance to LSP teaching and learning. The article concludes with the implications of this study for LSP practice and theory, and its wider implications for language teaching and learning in general.

2. Theoretical framework

2.1. Video in the language classroom

Video-based methodologies are well-established in second language teaching. According to Goldstein and Driver (2015: 1), the earliest paper on the subject dates back to 1947 and was an article by J.E. Travis on “The Use of the Film in Language Teaching and Learning”. In 1983, Willis established key roles for video in the classroom such as language focus, skills practice, stimulus and resource material (Willis, 1983: 29-42) and during the 1980s and 1990s, a vast quantity of video materials were specifically developed for use in the foreign language classroom, and language methodologists encouraged teachers to integrate video into foreign language teaching (Allan, 1985; Cooper, Lavery & Rinvoluceri, 1991). However, during the 1980s and 1990s, video was largely used as a static resource with classroom activities centred around viewing and listening to the video, or teaching the culture of the target language (Gardner, 1994; Nikitina, 2010). Video was often seen as a type of reward or light relief, often shown on a Friday afternoon or at the end of term.

In recent years, advances in digital technology have created exciting opportunities for using video in language teaching and learning. Video digital technology has made it easier to produce and edit video in a classroom setting as it is highly accessible with much of the technology already existing on students’ mobile phones, ipods and ipads. On the internet, video editing software such as *Windows Movie Maker* can be

downloaded for free and students can edit their videos easily. Research on video production as a tool for language learning and teaching has thus started to emerge with researchers examining the potential of digital video creation as a tool to enhance language learning. (Dal, 2010; Goldstein & Driver, 2015; Hafner and Miller, 2011; Shrosbee, 2008). Several case studies have been carried out in which researchers evaluated the effectiveness of video-making projects conducted in their own language classrooms (Goulah, 2007; Gromik, 2012; Kearney, Jones and Roberts, 2012; Naqvi & Mahrooqi, 2016; Nikitina, 2010; Reyes, Pich and Garcia, 2012). Between 2008 and 2010, the European funded *Divis* project (*Digital video streaming and multilingualism*) also aimed to encourage, motivate and equip language teachers to include video production in their teaching (1). The abovementioned studies demonstrate that digital video creation is not a new idea, and indicate that it is becoming an increasingly popular practice amongst researchers and teachers. However, it is also clear that very little research has been conducted on the integration of digital video creation in language teaching, and even less on its implications for developing language skills and other skills such as critical thinking, social and collaborative skills (Naqvi & Mahrooqi, 2016: 51). Caws and Heift (2016: 129) further argue that “the current culture of CALL, and, more specifically, the growing role of digital media in the daily life of learners, cannot be ignored.” In particular, digital video creation in the LSP context is entirely unexplored and it is thus timely to examine its integration in this area.

2.2. Languages for specific purposes

Languages for Specific Purposes (LSP) is a term with many definitions, interpretations and applications. Sager, Dungsworth and McDonald (1980: 68) define it as “specialist-to-specialist” communication, but this definition does not necessarily include the situation of the language learner who may not yet be a specialist in their domain. Chambers, (1996: 233) emphasises the need to take into account that different levels of specialisation may exist amongst learners, that language learners “may initially be non-specialists both in the language and in the subject they are studying.” Dudley-Evans and St. John (1998: 4-5) provide a detailed definition of English for Specific Purposes (ESP) arguing that ESP is designed to meet specific needs of the learner, using the underlying methodology and activities of the disciplines it serves and is centred on the language, skills, discourse and genres appropriate to these activities. This definition highlights the core concepts of LSP, that it is driven by the need to respond to students’ specific linguistic needs and uses the methodologies and activities needed to help learners enter the discourse community of the relevant discipline. This view of LSP is consolidated by Arnó-Macià (2014: 5) who argues that “since LSP teaching aims at helping students enter particular discourse communities, its methodology draws on relevant activities and practices”. For the purposes of this study, we will use this definition of LSP as a form of language teaching driven by students’ specific linguistic needs. In this case, the participants in our study are second-year undergraduate students on the BSc in Biotechnology programme and are thus non-specialists both in the French language and in the field of Biotechnology. Their language programme uses specific methodologies and activities which aim to help them enter the discourse community of Biotechnology.

Swales (2000: 59) traces research in LSP back to the 1960s when Halliday, McIntosh and Strevens (1964) highlighted the lack of investigation into the specialised material required to teach English to groups with specific linguistic needs such as power station engineers in India or police inspectors in Nigeria. However, Gollin-Kies, Hall and Moore (2015: 18) cite several examples of specialised goal-oriented courses prior to the development of LSP as a self-identified field such as a 1932 book designed to teach medical Arabic for medical workers in Syria and Palestine, and the introduction of German as a Foreign Language into the curriculum of a medical school in Shanghai, China in 1907. Early studies in LSP tended to be largely quantitative lexicostatistical studies providing information on specialist terminology and on which syntactic structures occurred most frequently in scientific prose (Chambers, 1996: 233). Swales (2000: 59) describes this early LSP research as descriptive and “basically textual or transcriptal”. However, over the years, challenges to this descriptive, textual tradition of work in LSP have arisen. There have been challenges to the simplistic relationship between linguistic analysis and classroom activities (Widdowson, 1998; Hutchinson and

Waters, 1987) together with new influences on LSP such as the development of the communicative approach to language learning (Chambers, 1996), the use of corpus linguistics data in LSP courses (Rodgers, Chambers and Le-Baron, 2011) and content and language integrated learning (CLIL) (Dalton-Puffer, Nikula and Smit, 2010: 1).

In more recent years, the field of LSP has been further shaped by factors such as increasing globalisation and the development of new communication technologies. (Gollin-Kies et al., 2015: 29-33). Globalisation has led to an increased demand for the teaching of foreign languages for specific purposes (Gollin-Kies et al., 2015: 35; Uber Grosse & Voght, 2012: 191) and one of the challenges of LSP teaching is to prepare students for “globalized academic and professional contexts” (Arnó-Macià, 2014: 15). Recent studies have also highlighted that advances in technology have revolutionised LSP language education, (Uber Grosse & Voght, 2012: 191) and that there is a need for special attention to be paid to LSP within the context of the integration of technology into language education (Arnó-Macià, 2012: 89). It is generally agreed that technology has transformed LSP teaching and learning in a number of ways. The role of IT in different areas of LSP research (Arnó, Soler & Rueda, 2006) and the design and implementation of online LSP materials (González-Pueyo, Foz, Jaime & Luzón, 2009) have been studied. Researchers acknowledge that developments in CALL, applied linguistics and the pervasive use of technology in communication have revolutionised LSP teaching (Arnó-Macià, 2012: 89). Uber Grosse and Voght (2012: 191) underline that technology gives LSP learners “instant access to current information about target languages and cultures” and that the Internet has “made it possible for LSP teachers and learners to access instantly rich resources of authentic language materials in their content field”. García Laborda (2011: 106) also highlights that because of the internet, “LSP materials that were difficult to find until recently (...) are now readily accessible and usually free”. Similarly, Arnó-Macià (2012: 89) outlines the ways in which emerging technologies have been integrated into the LSP classroom:

Through technology, LSP teachers and researchers can access discipline-specific materials and situations and compile corpora of specialized texts. Computer-mediated communication provides learning tools and a gateway to the discourse community. Technology also provides opportunities for collaborating, creating virtual environments and online courses, and fostering learner autonomy.

More recently Bárcena, Read and Arús' (2014) edited volume looks at LSP in the digital era and examines the impact of developments in the use of technologies such as CALL, wikis, corpus-based approaches and natural language processing on LSP, while the Gollin-Kies et al. (2015) volume also looks at the impact of new technologies on LSP teaching and learning in the 21st century in their volume on LSP.

However, while it is evident that the role of technology in LSP teaching and learning has been both examined and advocated, the role of video and more particularly video creation has not been explored within the LSP context.

2.2. Video creation in LSP

While the area of video creation in LSP has been hitherto unexplored, research in the related fields of LSP and video in the language classroom in general, point to video production as a particularly appropriate teaching and learning tool in LSP. LSP is traditionally a multidisciplinary activity which requires the learner to engage not just with the target language but also with disciplinary knowledge. Digital video creation enables language learners to link the learning of the target language with the learning of other content linked to their discipline. It further enables them to do this within a realistic context, reinforcing the principle that tasks for LSP learners should be as realistic for the learners' language goals as possible (García Laborda, 2011: 104) and use 'real-world' language in 'real-life' situations (Secules, Herron & Tomasello, 1992). LSP learners can thus blend language learning with disciplinary learning in a 'real-world' context through video production.

LSP teaching must also move with developments in new technologies as it is vital that “LSP methodologies should be rooted in how technology is used in real-life professional practices” (Arnó-Macià, 2014: 15-16). Digital video production gives learners the

opportunity to embrace not just the practice of creating a video but also the multimodality of the target language. The rise in multimodality is a particularly striking trend in technologically-mediated communication as the development of new communication technologies has enabled LSP teachers and students to download and engage with large amounts of multimodal data which provide excellent opportunities for language learning (Gollin-Kies et al., 2015: 43). Digital media carry words, sounds and images and enable learners to embrace the multimodality of language. It can be argued that language has always been multimodal and it has always been “a mixture of sound, words, images created in the mind, and gestures used in contexts full of objects, sounds, actions and interactions” (Gee & Hayes, 2011: 1). Multimodality is equally considered to be a defining characteristic of CALL (Guichon and Cohen, 2016: 509). For the LSP learner, it is particularly important to keep pace with the rise in multimodality afforded by digital media as it is a central tenet of LSP that their language learning must be both contextualised and authentic.

Video creation also assists the LSP learner in the acquisition of a wider range of professional and social skills. Arnó-Macià (2014: 9) draws attention to the centrality of social and critical skills for the LSP learner, arguing that LSP courses play a vital role in the integration of professional communication skills with key social and critical competences that students need to participate in society. It is particularly important that LSP learners acquire those communication skills necessary to participate in 21st century society. Goldstein & Driver (2015: 117) cite the acquisition of ‘21st century skills’ as amongst the goals of any digital video creation project:

The primary goals are situating language through practical engagement in the creation of digital artefacts. This is achieved through the process of guided reflection, critical thinking, performance, debate, design, creativity and other competences often referred to as ‘21st century skills’.

Video production enables LSP learners to think critically about the topic they have chosen to present, to express their ideas and opinions, to debate, to perform and above all to be creative. It gives learners choices, not only about what to say, but also how to say it and how to present a point of view (Dal, 2010: 5). The development of these skills is vital for the LSP learner, thus rendering video production a very appropriate tool for language learning in this domain.

The task-based nature of digital video creation is equally advantageous for the LSP learner. Video production is very much a learner-centred, practical, hands-on, creative project. It is essentially a form of task-based learning which embraces the social constructivist view of constructing knowledge and meaning in a social context through practice (Arnó-Macià, 2014: 14; Goldstein et al., 2015: 118). Nikitina (2010: 22) argues video-making projects include all the core elements of progressive language pedagogy.

(...) involving language learners in the production of digital video in the target language follows constructivist perspectives on teaching and learning since the main tenets of progressive language pedagogy, such as learner-centeredness, activity-based learning, and a communicative approach, put emphasis on the active involvement of the learners in the teaching/learning process and call for collaboration between learners. All these elements are present in the video-making activity.

Through video creation LSP learners learn to negotiate meaning through the creation of a digital artefact. Students become ‘producers’ of language (Dal, 2010: 3; Shrosbee, 2008: 75). This is vital in language learning as every human is both a producer and a consumer of language and digital media enable learners to be both producers and consumers of language (Gee and Hayes 2011: 2-3). By producing videos on subject areas relevant to their discipline, they produce language, negotiate meaning, communicate and collaborate and thus engage in a language learning activity which is both meaningful and pedagogically effective.

While research in the related fields of LSP and video in the language classroom point to video creation as a particularly appropriate tool for the LSP classroom, this study investigates student perceptions of this activity and reports on its effectiveness from a

learner perspective. It also analyses the videos created in order to see if the outcome of this teaching and learning activity substantiates the initial findings.

3. Current Study

This study is based on a digital video project conducted by second-year university students of French. It seeks firstly, to examine their perceptions of the effectiveness of video production to develop their language skills within the context of LSP and secondly, to determine if these findings are substantiated through an analysis of the videos created. The study thus attempts to add to the body of research on both video in language learning and LSP.

3.1. Participants

The students participating in this study were twenty-three second-year students in NUI Galway taking the BSc. in Biotechnology, a domain of Science which is often described as the application of biology for the benefit of humanity and the environment. Students on this four-year programme study a wide range of subjects such as biology, chemistry, biochemistry, microbiology, genetics, toxicology and pharmacology. They tend to find employment in industries such as biopharmaceuticals, diagnostics, healthcare and the environment. In addition to studying the relevant science subjects, students on this programme study either French or German for the first three years and tend to continue with the language they have previously studied at second-level. Many students choose this course because of its unique offering of Science and a language, and are aware of the prominent positions of France and Canada in the biotechnology industry.

Students have three hours of French per week and their programme aims to enable them to acquire the specialised French for Biotechnology they need to enter this discourse community. Activities such as text analysis, roleplays, simulations, communication games, grammar activities, project work, multimedia lab work, group discussions and presentations are all used to promote their engagement with the target language and develop their knowledge of French for Biotechnology. Classes and activities are generally based around contemporary topics of interest in the field of Biotechnology such as stem-cell research, gene therapy, cancer research, marine biotechnology and so on.

3.2. Project

The main objectives of the video creation project were to help students to further develop their language skills in French, to acquire a more in-depth knowledge of specialised French for Biotechnology, to develop the practical sub-skills of video production and editing and to acquire other key competences such as critical thinking, creativity and teamwork. Students were asked to create three to four minute videos in groups of three (2) on an area of contemporary Biotechnology research of their choice, describing the area involved and highlighting those aspects of it they found most interesting. The only instruction they were given was that all members of the group were to speak and feature in the video. They were subsequently taken for a preparatory class where guidance was given on how to create storyboards and edit the end product. It was explained to them that video production comprises three key phases; pre-production, production, and post-production. They learned that pre-production is primarily a planning phase encompassing storyboarding (defining each individual shot as visual representations), location scouting, scripting and audience identification. They learned that production is the creation of the footage required for the video in accordance with the storyboard and script while post-production centres on the editing required, both audio and video, to produce the final film. They were introduced to software such as *Windows Moviemaker* and *Me Move* and advice was given on the creation of storyboards and video production and editing. They were given six weeks to create their video outside of class time and they were asked to submit it as an MP4 file via the University's internal server *Blackboard*. They were not given any other instructions or restrictions in order to allow for maximum creativity. They were free to choose their own themes for the videos and their own formats. For example, they could make a short movie, a documentary, an interview, a promotional video, a debate or talk show. The videos were to be screened during class after the six week preparatory period so that they could all view each other's work. The teacher believed this would

motivate them to be more creative and produce a better video. In total, eight videos were produced by seven groups of three students and one group of two. The videos produced were based on a variety of topics including hybrid embryos, genetically modified foods, the Zika virus, animal testing in Biotechnology research and designer babies.

3.3. Methodology

The effectiveness of this teaching and learning activity was evaluated in a number of ways. At this point, it is imperative to remind ourselves of the complexity of the evaluation of CALL tasks, and that learning is in general an unquantifiable concept (Nokelainen, 2006: 183). In this evaluation, we are seeking to establish the effectiveness of digital video creation as a tool for these particular learners, echoing in many ways Chapelle (2001: 5), who argues that “an evaluation has to result in an argument indicating in what ways a particular CALL task is appropriate for particular learners at a given time.” The evaluation in this study includes the factors generally included in an evaluation in CALL: the actors (learners), the tool that is being utilised and the artefact created (Caws and Heift, 2016: 128). In this instance, the effectiveness of the activity from a student perspective was evaluated using a mixed method of data collection, combining student questionnaires and semi-structured group interviews. Part One of the questionnaire was designed to gather background information about the subjects’ language and video production skills by asking them how long they had been studying French, if they had ever created videos or used video editing software before, how difficult or easy they found it to use the editing software, what device they used for filming and what video editing software they used. Part Two of the questionnaire focussed on what they felt they had learned from creating the video. In questions six to eight, students were asked to describe how helpful or unhelpful they found the video creation activity for learning vocabulary and grammar, using a four-point Likert scale ranging from “very helpful”, “helpful”, “a little helpful” to “not at all helpful”. They were also asked to compare the use of video production with traditional ways of learning French. Question nine asked if they thought they had acquired any other skills (apart from language) from participating in this project while question ten was an open question in which they were asked to list what they felt were the main advantages and disadvantages of creating a video to learn French for Biotechnology. The final question asked students for any general comments or recommendations in relation to the use of video creation for learning French for Biotechnology.

The semi-structured group interview took place in a thirty minute session after class. The interview aimed to delve deeper into students’ experiences and invited them to reflect carefully on their experience of video creation for language learning. They were asked at the beginning of the session to be honest about their experience of the video creation project, even if the experience was negative. The questions were intended to enable students to express their ideas without any undue influence by the teacher conducting them. Students were simply asked what they liked or did not like about the video creation project and what they felt they had learned. As various points were raised, they were further probed by the interviewer and expanded upon by other students. The discussion was allowed to unfold in a natural way and the teacher did not comment on the points made. The session was recorded and subsequently transcribed.

Quantitative data from the questionnaire was subsequently examined, firstly to establish background information on students’ level of language learning and experience of video production and editing, and secondly to rate their perceptions of the usefulness of video creation to learn French. Qualitative data from the open questions in the questionnaire and the group interview were manually analysed and categorised into broad themes to report on key aspects of students’ perceptions of video creation as a language learning tool for LSP.

The videos were then analysed in order to see if the learners’ claims could be substantiated. It was not the purpose of this study to use pre- or post-tests to measure progress or vocabulary acquisition, but rather to see if the outcome of this activity was indicative of its pedagogical value. The videos were thus closely examined by both researchers in this study, and observations were recorded and subsequently compared with initial data stemming from the questionnaire and interview. We recognise that in

observational data of any description, there is a limitation which must be acknowledged, that the conclusions drawn must be taken as indicative as opposed to conclusive (Harbon and Shen, 2015: 463). The analysis of the videos created thus seeks to establish whether the artefacts created by students were *indicative* of the pedagogical effectiveness of the activity.

4. Results and Discussion

4.1. Questionnaire analysis: quantitative data

Twenty-two students in total completed the questionnaire. From Part One, it was clear that most students had been studying French for five to eight years. Most students had never created videos previously, with only six students stating they had made videos before, during a transition year of secondary school. Five of these six students had used video editing software before and indicated that they had used either *Microsoft Movie Maker* or *Video Star*. It was evident that video production and editing was a new activity for the vast majority of participants. When asked how easy or difficult they found it to use the editing software, responses indicated that in general, students found this aspect of the project challenging. This question was directed only to the seventeen participants who had never used video editing software before and the majority of students described it as "a little difficult". Eight videos in total were created by this group, two of which were filmed using android phones, two using iPhones, two with iPads and two using laptop cameras. While students were offered the use of Camileo clips, it was clear that they preferred to use their own devices, thus corroborating Burston's (2016: 5) view that mobile device ownership has reached a point where it is feasible to expect students to use their own devices for language learning. When asked what video editing software they used, the majority of students had used *Microsoft Movie Maker*, while other software used included *Filmauro Wondershare*, *Microsoft Video Editor*, *Video pod*, *Splice* and *NCH*. The analysis of Part One of the questionnaire, therefore, indicated that while students were approaching this project with at least 5 or more years of French behind them, most had no experience of video production or editing and found the editing aspect a little difficult, thus perhaps challenging the view that all members of the current generation are "digital natives" (Prensky & Heppell, 2008).

In Part Two, students were asked to describe how helpful (or unhelpful) they found this project for learning French and responses were extremely positive. When asked to what extent they found creating the video helped them acquire French vocabulary relating to Biotechnology, all students replied that they found it either "very helpful" or "helpful". When asked to what extent they found it helpful to improve their knowledge of French grammar and structures, twenty students felt that it was either "very helpful" or "helpful", while two described it as "a little helpful". Students were thus overwhelmingly positive in their evaluation of the usefulness of the project to improve their language skills, particularly in the domain of the acquisition of specialised vocabulary. When asked to compare how they found video production as a tool to learn French compared to traditional ways they were taught in the past, all but one student described it as either "a lot more helpful" or "more helpful".

In Section 2.3, it was highlighted that research in the related fields of video production in language learning and LSP point to video production as a particularly appropriate tool for LSP. The quantitative data elicited from this analysis is indicative of an overwhelmingly positive response to the usefulness of video creation for language learning in the LSP domain. Students perceive it as a helpful means of improving their knowledge of French grammar and structures and find it particularly helpful to acquire specialised vocabulary relating to Biotechnology. One of the core concepts of LSP underlined in Section 2.2, is that it is driven by the need to respond to students' *specific* linguistic needs and that it should use activities to help learners to enter the discourse community of their relevant discipline (Arnó-Macià: 2014: 5). This project thus enables them to focus on the terminology and structures most relevant to their field of study. In addition, it compares very favourably with traditional ways of learning French. It is however, the comments of the learners, both in the questionnaires and the semi-structured group interview, which shed more light on the pedagogical potential of video creation in an LSP context.

4.2. Qualitative data from the questionnaire and semi-structured group interview

Qualitative data was elicited from four open questions in the student questionnaire. When asked in question nine to compare video production with traditional ways of learning French, participants were asked to explain their answer. Question ten asked them what other skills (if any) apart from language, they felt they had acquired from participating in the project while Question eleven asked them to list, in their opinion, the main advantages and disadvantages of creating a video to learn French for Biotechnology. The final question invited them to make any further comments or recommendations with regard to the video creation project.

As mentioned in the previous section, twenty-one out of the twenty-two students surveyed described video production in Question nine as more helpful when asked to compare it to traditional ways they were taught in the past. When asked to explain their answer, eleven students focussed on how this project made them more aware of their pronunciation and accent in French, explaining that because they had to listen to themselves speaking French, that they became more conscious of pronunciation errors and tried to correct them, very often by re-filming segments. This self-awareness of their speaking skills in French together with the opportunities it afforded for self-correction is mentioned several times in other open questions and appears to be a key factor in their perception of the usefulness of this project. Seven students also mentioned that it provided them with variety in their language learning and that it was "fun".

In Question ten, students elaborated on other skills they felt they had acquired from participating in the project. Teamwork seemed to be a key skill they thought they had developed with sixteen students referring to it as a skill acquired. Ten students also referred to the development of their teamwork skills during the interview and elaborated that this was the first time during their time at University that they had really had to work together in this way. They explained that tasks cannot really be divided up when making a video as all team members have to participate actively in each stage of the process, corroborating Nikitina's (2010: 22) argument that video-making requires learners to collaborate and work together. Eleven students also mentioned organisational skills in the questionnaire as an important feature, with some students further explaining that the project required them to manage their time, work to a deadline, organise schedules, equipment and so on. Ten students listed video production and editing skills as another key competency acquired, a point that was echoed in the group interviews by twelve students. In the group interview, these students observed that they had never done anything like this project before and that the acquisition of the technical skills necessary to produce and edit a video was something they really valued. Communication skills were also mentioned in the questionnaire by five of the students surveyed. These reactions corroborate the view that video creation enables students to acquire key competencies or "21st century skills" (Goldstein & Driver, 2015: 117). The findings show that students felt that this project enabled them to develop organisational, communication, technical and teamwork skills and thus also confirm Arnó-Macià's (2014: 9) argument that LSP courses have a key role to play in the integration of professional and social communication skills.

Question eleven asked the students to list up to three advantages and disadvantages of creating a video to learn French for Biotechnology. Fifty-four advantages in total were listed (an average of 2.5 per student) while twenty-nine disadvantages were listed (an average of 1.3 per student). The main advantage perceived by students and specified in the questionnaire, was that video creation enabled them to see and listen to themselves speaking French, thus allowing them to improve their pronunciation, accent and spoken French in general, with seventeen students citing this as an advantage. Likewise, in the group interview fourteen described the main advantage of creating video as the capacity to hear and see themselves speaking French, and to be able to correct their own errors by re-filming when necessary. On the whole, the responses suggest that self-awareness of their spoken competencies in the target language together with the ability to self-correct was perceived by students as the key advantage. A substantial number of students (nine) also mentioned in the questionnaire, that creating the video had allowed

them to practice their speaking skills and that they had gained confidence in speaking French, while a further ten students described how they appreciated the opportunity this project gave them to learn in a different way. The other key advantages mentioned were the teamwork dimension to the project, the opportunity it gave them to be creative and several students described it as a “fun” way to learn. These findings were echoed in the group interview where students expanded on the “fun” aspect to the project. Ten students used the word “fun” in the group interview and explained that they really enjoyed completing this project and had laughed a lot in the process. Eight students spoke about the opportunity to be creative, explaining that they appreciated the freedom to pick their own topic and presentation format. Creativity and activity-based learning are core elements of current language pedagogy and students’ reactions confirm their willingness to become creative learners and ‘producers’ of language (Dal, 2010: 3; Shrosbee, 2008: 75). Four students mentioned the opportunity to explore scientific topics in an in-depth way in French as an advantage in the questionnaire and this point was reiterated by five students in the group interview. Learners thus also appreciated the opportunity to link their knowledge of the target language and their discipline, to use language in a realistic context (García Laborda, 2011: 104).

In terms of disadvantages, there appeared to be two main elements the students found difficult. In question eleven of the questionnaire, ten students commented on the time-consuming nature of the project, while a further ten described how difficult they found the editing process. Other disadvantages mentioned included the difficulty of selecting a topic, working in a group and finding times to suit everybody. These findings were echoed in the group interview with twelve students referring to the complexity of the editing process as a disadvantage and ten describing how time-consuming the project was. When probed as to how much time the project took to complete, responses ranged from six hours to sixteen and students described how difficult it is to create storyboards, time segments, prepare scripts and edit. The difficulties encountered by students on the technical side of video creation were largely unanticipated in this study as it appears to defy an underlying assumption in CALL, MALL and CMC research, that students are part of a digital generation and entirely capable of engaging with the necessary technology. Instead, data from this study highlights how difficult they can find it and warns us as researchers to refrain from too many assumptions in this domain.

The final question on the questionnaire invited students to make comments or recommendations in relation to the use of video creation for learning French for Biotechnology. Eight students had no comments and of the remaining fourteen, five commented that they would have liked more training on video editing prior to commencing the project. The remaining nine mostly reiterated that they had enjoyed the project and that they would recommend it be used with future groups of students.

4.3. *Observational data from the analysis of the videos created*

Observational data from the analysis of these videos indicates that the key claims made by students in the questionnaires and group interviews; including improved language skills, acquisition of new technical skills, and improved organisational, communication, creative and teamwork skills can be substantiated.

In terms of their language skills, students had claimed that this activity had enabled them to improve their language skills, particularly their acquisition of specialised vocabulary relating to Biotechnology. Their acquisition and successful use of specialised vocabulary was evident in all eight videos. In some instances, they were consolidating the use of vocabulary acquired in class, and in others they demonstrated their acquisition of terminology previously unknown as indicated in the following three examples. Video three was based on the topic of hybrid embryos and here specialised terms were used such as “la pénurie d’ovocytes humains” (*the shortage of human embryos*), “des lignées de cellules souches” (*stem cell lines*), “l’ADN” (*DNA*), “greffer” (*to transplant*) and “des défauts génétiques” (*genetic defects*). Video four explored the topic of the Zika virus and over the course of the video, the learners demonstrated their acquisition of specialised terminology such as “les maladies parasitaires” (*parasitic diseases*), “des moustiques génétiquement modifiés” (*genetically modified mosquitos*), “l’accouplement” (*mating*) and “des vecteurs” (*vectors*). Video six on the subject of genetically modified foods showed students using terms such as “l’hôte” (*the host*),

"une carence" (*a deficiency*), "des insects ravageurs" (*devastating insects*) and "le génie génétique" (*genetic engineering*).

Participants had also highlighted the usefulness of this project to improve their pronunciation, accent and general speaking skills and had particularly emphasised the opportunity digital video creation gave them to see and hear themselves, and self-correct before submitting the final product. While it is not within the scope of this study to measure the exact level of improvement in pronunciation and speaking skills as a result of producing these videos, it was evident from a close analysis of the videos that the participants' displayed a high level of accuracy in pronunciation, intonation and accent. It was also apparent that a considerable amount of editing had taken place in all eight videos (see below), thus confirming the participants' claims that they had re-recorded on several occasions in an effort to correct their pronunciation.

Teamwork, organisational, communication and technical video production skills were all identified as key competencies acquired during the course of this project. Each of the eight videos portrayed an obvious style which was adhered to throughout the video; amongst them a news report, a college debate, a chat show discussion, and a scientific report. Each group maintained the style of their individual video by sourcing appropriate locations and costumes, and by scripting the video to use appropriate language; thereby evidencing the level of teamwork, communication and planning involved. Obvious examples include the multiple locations that were used in the filming of five of the videos. For example, in video two, based on the development of a device for astronauts to analyse their own blood samples, five different locations were used including an interview room, hospital entrance, hospital bed, office setting and foyer. Similarly in video one about designer babies; four locations were used including a kitchen, laboratory, corridor and classroom. Video seven on ethical issues in Biotechnology used a classroom setting with a teacher and two students to explore the topic, while videos five and eight used a debate amongst speakers to frame the topic. Videos three and four used news report settings with multiple locations to present their topic.

Technical skills acquired by all eight groups during video production include the framing of the shot; allowing adequate head room and looking room, shot composition; close-ups (CUs) medium close-ups (MCUs), mid-shots (MS), medium long shots (MLS), establishing shots and long shots (LS); panning, tilting, zooming. In keeping with the scientific nature of the subject, the groups framed many shots as a piece-to-camera (PTC), a key technique used on news reports or factual programmes, where the presenter would introduce a topic by speaking directly to camera. Evidence of technical skills during post-production included the use of opening titles and credits by seven of the eight groups, with five of these groups also including a soundtrack (four groups sourced French language songs), with the eighth group also applying credits during post-production. Dissolves and fades were used in four of the videos to move between shots, while cutaways (the use of footage or still images appropriate to the matter being discussed) were used in three videos.

Creativity, while subjective, was clearly evidenced in all videos by the participants acting out roles they had assigned to themselves; news reporters, presenters, teachers, victims of the Zika virus, chefs, hybrid organisms, astronauts, patients, researchers, scientists in Hazmat suits, doctors, babies and many more. They dressed up in character and used props, music and humour to present their topics creatively and in two videos used out-takes at the end of the video to show humorous moments during filming. Also, on a practical level, because all student members of each team had to feature and speak in their video this required them to take turns filming and plan each video carefully, thereby improving communication and teamwork.

5. Conclusion

The principal aim of this project was to determine the effectiveness of digital video creation as a teaching and learning tool in the LSP context through an investigation of student perceptions of the usefulness of this activity and a subsequent comparison of this data with an analysis of the digital artefacts created. The quantitative and qualitative data gathered indicate an overwhelmingly positive response to the use of this tool in LSP. The participants in this study found it to be a very helpful means of

improving their language skills, especially in the domain of the acquisition of specialised vocabulary. In particular, the usefulness of this project to improve their pronunciation, accent and general speaking skills was highlighted and participants explained that digital video creation gives learners the unique opportunity to see and hear themselves and to self-correct before submitting the final product. Teamwork, organisational, communication and video production skills were all identified as key competencies acquired during the course of this project, thus demonstrating that video creation can play a key role in the acquisition of professional and social skills, a factor identified as a key tenet of LSP courses. "Fun" was a frequent term used by students in feedback gathered, and responses showed that students appreciated the opportunities to be creative and to engage in task-based learning that this project gave them. Students thus perceive digital video creation as more than just a means to improve their language skills, but also as a means to acquire other key social and professional skills in a creative and fun way. The analysis of the videos indicated that these findings could be substantiated, and the high quality of the videos produced demonstrated that engaging in digital video production had had a strong impact on these learners.

Going forward, however, the data gathered shows the students must be given more support and help with the production and editing processes. Video production and editing were new activities for the vast majority of these learners and a greater level of technical assistance will have to be provided in the future to make this a better learning experience. While 21st century students may be considered "digital natives" (Prensky & Heppell, 2008), this does not mean that they can master all aspects of digital technology without assistance or guidance, and responses from students indicate that this project was more challenging and time-consuming from a technical perspective than anticipated. If anything, this study strengthens Prensky's view that our young people need to become digitally wise in order to be able to keep up in what he terms "an unimaginably complex future" (2012: 182).

In the broader context of language teaching and learning research, this study highlights the pedagogical potential of digital video creation in the language classroom. It corroborates studies which point to video creation as a pedagogically useful tool for language learning and teaching and extends this largely unexplored area of research to the field of LSP. While LSP research advocates the need to integrate emerging technologies into the 21st century LSP classroom, this study gives a very practical example of how this can be done. It demonstrates that digital video creation embraces many of the core elements underpinning LSP pedagogy by enabling language learners to link their language learning with their discipline of study and to do so in a 'real-world' or 'real-life' situation. It also gives them the opportunity to keep pace with the multimodality afforded by digital media and thus means their language learning is both contextualised and authentic. In addition, it assists LSP learners in the acquisition of those professional, social and communication skills necessary to participate in 21st century society. Critical thinking, creativity, performance and autonomy are all skills developed through digital video creation. This study thus contributes to the area of LSP research and to the broader areas of digital technology in language learning by demonstrating that digital video creation is a pedagogically beneficial and meaningful activity for LSP learners.

References

- Allan, M. (1985). *Teaching English with video*. Essex: Longman.
- Arnó-Macià, E. (2012). The role of technology in teaching languages for specific purposes courses. *The Modern Language Journal, Focus Issue: Languages for Specific Purposes*, (96): 89-104.
- Arnó-Macià, E. (2014). Information technology and languages for specific purposes in the EHEA: options and challenges for the knowledge society. In Bárcena, E., Read, T. and Arus, J. (eds.) *Languages for specific purposes in the digital era*. Heidelberg; New York; Dordrecht; London: Springer.
- Arnó, E., Soler, A. and Rueda, C. (eds.) (2006). *Information technology in languages for specific purposes*. New York: Springer.

Bárcena, E., Read, T. and Arus, J. (eds.) (2014). *Languages for specific purposes in the digital era*. Heidelberg; New York; Dordrecht; London: Springer.

Burston, J. (2016). The future of foreign language instructional technology: BYOD MALL. *The EUROCALL Review*, 24(1): 3-9.

Caws, C. and Heift, T. (2016). Evaluation in CALL. Tools, interactions, outcomes. In Farr, F. and Murray, M. (Eds.). *The Routledge Handbook of Language Learning and Technology*. London; New York: Routledge, 127-140.

Chambers, A. (1996). LSP theory and second language acquisition. In Hickey, T. and Williams J. (eds.), *Language, education and society*. Bristol: Multilingual Matters, 232-238.

Chapelle, C. (2001). *Computer Applications in Second Language Acquisition: Foundations for Teaching, Testing and Research*. Cambridge: Cambridge University Press.

Cooper, R., Lavery, M. and Rinvolutri, M. (1991). *Video*. Oxford: Oxford University Press.

Dal, M. (2010). Digital video production and task-based language learning. *Ráostefnurit Netlu-Menntakvika*. <http://netla.khi.is/menntakvika2010/alm.021.pdf> (accessed June 20, 2017)

Dalton-Puffer, C., Nikula, T. and Smit, U. (2010). *Language use and language learning in CLIL classrooms*. Philadelphia; Amsterdam: John Benjamins.

Dudley-Evans, T. and St. John, M.J. (1998). *Developments in ESP. A multi-disciplinary approach*. Cambridge: Cambridge University Press.

Dugartsyrenova, V. and Sardegna, V. (2016). Developing oral proficiency with VoiceThread: Learners' strategic uses and view. *ReCALL*, 29(1): 59-79.

Duman, G., Orhon, G. and Gedik, N. (2014). Research trends in mobile assisted language learning from 2000 to 2012. *ReCALL*, 27(2): 197-216.

García Laborda, J. (2011). Revisiting materials for teaching languages for specific purposes. *The Southeast Asian Journal of English Language Studies*, 17(1): 102-112.

Gardner, D. (1994). Student-produced video documentary: Hong-Kong as a self-access resource. *Hong Kong Papers in Linguistics and Language Teaching*, 17: 45-53.

Gee, J.P. and Hayes, E.R. (2011). *Language and learning in the digital age*. New York: Routledge.

Goldstein, B. and Driver, P. (2015). *Language learning with digital video*. Cambridge: Cambridge University Press.

Gollin-Kies, S., Hall, D. and Moore, S.H. (2015). *Language for specific purposes*. New York: Palgrave Macmillan.

González-Pueyo, I., Foz, C. Jaime, M. and Luzón, M.J. (eds.) (2009). *Teaching academic and professional English online*. Bern: Lang.

Goulah, J. (2007). Village voices, global visions: digital video as a transformative foreign language tool. *Foreign Language Annals*, 40(1): 62-78.

Gromik, N. (2012). Cell phone video recording feature as a language learning tool: A case study. *Computers & Education*, 58: 223-230.

Guichon, N. and Cohen, C. (2016). Multimodality and CALL. In Farr, F. and Murray, M. (Eds.). *The Routledge Handbook of Language Learning and Technology*. London; New York: Routledge, 509-521.

Hafner, C. and Miller, L. (2011). Fostering learner autonomy in English for science: a collaborative digital video project in a technological learning environment. *Language Learning & Technology*, 15(3): 68-86.

Halliday, M.A.K., Strevens, P. and McIntosh, A. (1964). *The linguistic sciences and language teaching*. London: Longman.

- Harbon, L. and Shen, H. (2015). Researching language classrooms. In Paltridge, B. and Phakiti, A. (eds.). *Research Methods in Applied Linguistics. A Practical Resource*. London; New Delhi; New York; Sydney: Bloomsbury, 457-470.
- Hutchinson, T. and Waters, A. (1987). *English for specific purposes*. London: Longman.
- Kearney, M., Jones, G. and Roberts, L. (2012). An emerging learning design for student-generated 'iVideos'. *Teaching English with Technology. Special Issue on LAMS and Learning Design*. 12(2): 103-120.
- Naqvi, S. and Mahrooqi, R. (2016). ICT and language learning: A case study on student-created digital video projects. *Journal of Cases on Information Technology*, 18(1): 49-64.
- Nikitina, L. (2010). Video-making in the foreign language classroom: applying principles of constructivist pedagogy. *Electronic Journal of Foreign Language Teaching*, 7(1): 21-31.
- Nokelainen, P. (2006). An empirical assessment of pedagogical usability criteria for digital learning material with elementary school students. *Educational Technology and Society*, 9(2): 178-197.
- Prezky, M. (2012). *From digital natives to digital wisdom: hopeful essays for 21st century learning*. California, London, New Delhi: Sage.
- Prezky, M. and Heppell, S. (2008). *Teaching digital natives: partnering for real learning*. Thousand Oaks, California: Corwin Press.
- Prezky, M. (2001). Digital natives, digital immigrants. *On the Horizon* (MCB University Press). 9(5): 1-6.
- Reyes, A., Pich, E. and Garcia, M.D. (2012). Digital storytelling as a pedagogical tool within a didactic sequence in foreign language teaching. *Digital Education Review*, 22: 1-18.
- Rodgers, O., Chambers, A. and Le-Baron Earle, F. (2011). Corpora in the LSP classroom: a learner-centred corpus of French for biotechnologists. *International Journal of Corpus Linguistics. Applying Corpus Linguistics*, 16(3): 391-411.
- Sager, J.C., Dungworth, D. and McDonald, P.M. (1980). *English special languages: principles and practice in science and technology*. Wiesbaden: Brandstetter Verlag.
- Secules, T., Herron, C. and Tomasello, M. (1992). The effect of video context on foreign language learning. *The Modern Language Journal*, 76: 480-490.
- Shrosbee, M. (2008). Digital video in the language classroom. *JALT CALL Journal*, 4(1): 75-84.
- Swales, J.M. (2000). Languages for specific purposes. *Annual Review of Applied Linguistics*, 20: 59-76.
- Über-Grosse, C. and Voght, G.M. (2012). The continuing evolution of languages for specific purposes. *The Modern Language Journal, Focus Issue: Languages for Specific Purposes* 96: 190-202.
- Widdowson, H.G. (1998). Communication and community: the pragmatics of ESP. *English for Specific Purposes*. 17: 3-14.
- Willis, J. (1983). Implications for the exploitation of video in the EFL classroom. In McGovern, J. (ed.), *Video applications in English language teaching, ELT documents 114*. London: Pergamon Press, 29-42.

Appendix A

Questionnaire – Quantitative Data

Q1 How long have you been studying French?	3-5 years	0
	5-8 years	20
	More than 8 years	2
Q2 Have you ever created videos before this Semester?	Yes	6
	No	16
	Don't know	0
Q3 Have you ever used video editing software before this project?	Yes	5
	No	17
	Don't know	0
Q4 If you answered no to the previous question, how difficult or easy did you find it to learn to use the editing software?	Very easy	1
	Easy	2
	A little difficult	12
	Very difficult	2
Q5 What device did you use to film the video?	Android phone	2
	iPhone	2
	iPad	2
	Laptop camera	2
Q6 What video editing software did you use?	Windows Movie Maker	11
	NCH	2
	Filmora Wondershare	3
	VideoPad	3
	Splice	3
Q7 During the semester, you created a video on an area of Biotechnology of your choice. To what extent do you feel that creating this video helped you acquire French vocabulary relating to Biotechnology?	Very helpful	14
	Helpful	8
	A little helpful	0
	Not at all helpful	0
Q8 To what extent do you feel the creation of this video helped you to improve your knowledge of French grammar and structures?	Very helpful	7
	Helpful	13
	A little helpful	2
	Not at all helpful	0
Q9 How did you find video production as a means to learn French compared to traditional ways you were taught in the past?	A lot more helpful	12
	More helpful	9
	A little less helpful	1
	A lot less helpful	0

[1] www.divisproject.eu

[2] As there were twenty-three students in the group, we had seven groups of three and one group of two.

Literature review

Speaking Practice Outside the Classroom: A Literature Review of Asynchronous Multimedia-based Oral Communication in Language Learning

Eric H. Young* and Rick E. West**
Brigham Young University, USA

*ey390x@gmail.com | ** rickwest@byu.edu

Abstract

Classroom instruction provides a limited amount of quality speaking practice for language learners. Asynchronous multimedia-based oral communication is one way to provide learners with quality speaking practice outside of class. Asynchronous multimedia-based oral communication helps learners develop presentational speaking skills and raise their linguistic self-awareness. Twenty-two peer-reviewed journal articles studying the use of asynchronous multimedia-based oral communication in language learning were reviewed, (1) to explore how asynchronous oral communication has been used to improve learner speaking skills, and (2) to investigate what methodologies are commonly used to measure and analyze language gains from using asynchronous multimedia-based oral communication to improve learner speaking skills. In this study we present three principal findings from the literature. First, asynchronous multimedia-based oral communication has been used in conjunction with a variety of instructional methods to promote language gains in terms of fluency, accuracy and pronunciation. Second, the methods found in this review were technical training, preparatory activities, project-based learning, and self-evaluation with revision activities. Third, the majority of previous studies demonstrating the effectiveness of these methods have relied on learner perceptions of language gains rather than on recordings of learner speech.

Keywords: Oral, online, asynchronous, video, audio, language learning.

1. Introduction

In order for foreign language learners to succeed, they need a large quantity of high quality language practice. Although Clifford described time on task, or quantity, as “the primary determiner of language acquisition” (2002), it has also been described as “a necessary, but not sufficient, condition for learning” (Karweit, 1984: 33). Hirotani and Lyddon (2013) argued that quality of practice, exemplified in their study by an awareness-raising activity, is an important factor in the language learning.

Media-based oral communication can increase the quantity and improve the quality of language practice by providing more opportunities for speaking and more opportunities to raise learner awareness. Multimedia-based oral communication includes a variety of communication types, such as video conferencing through Skype, posting vlogs on YouTube, and turn-based video conversations using a voiceboard. Lin (2015) lauded the affordances of oral computer-mediated communication (CMC: an important type of multimedia-based oral communication) in his meta-analysis, stating that the “features of CMC seem to provide opportunities to create a social interaction context with more flexibility that cannot be afforded in a traditional face-to-face environment” (p. 262). Here it is useful to recall Clark’s (1994) criticism of many media-related studies, that media itself does not influence learning. Rather it is the instructional method that influences learning. Referring to his previous studies, Clark summarized his argument, stating, “any necessary teaching method could be designed into a variety of media presentations (p. 22). On the other hand, however, it is important to note that certain

media and technologies provide *affordances* that may not be otherwise available or that are more effectively used with those media and technologies.

In his book on distance and blended (a.k.a., hybrid) learning, Graham (2006) stated that online learning environments provide learners with flexibility in communicating outside the classroom. By communicating online, learners may increase their opportunities for speaking practice. Additionally, the digital nature of online communication makes it easier for learners to record and review their speech, allowing them to develop linguistic self-awareness. Both the opportunities and self-awareness promote increased speaking proficiency. Figure 1 illustrates these affordances and their relationship.

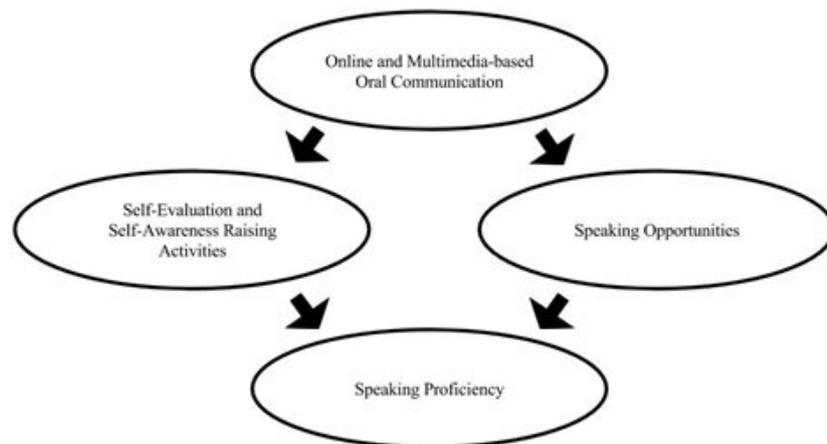


Figure 1. Relationship of online and multimedia-based communication to speaking proficiency.

Lin (2015) discussed these affordances in his meta-analysis of CMC use. Although he referred specifically to text-based communication, the affordances also apply to oral communication. He stated that CMC “provides L2 learners with an environment to practice language production at a reduced rate. The relatively reduced rate of exchange and lag-time induced by the text-chat software allows L2 learners ‘more time to both process incoming messages and produce and monitor their output’ (Sauro & Smith, 2010: 557)” (Lin, 2015: 264).

Similarly, in her meta-analysis of 14 studies involving CMC, Ziegler (2016) argued that CMC use provides learners with an opportunity to “notice [the] gaps between their interlanguage and the target language” (p. 575). Because of the time lag that Lin (2015) referred to, Ziegler (2016) found that CMC may be more beneficial to language learning than face to face communication in the target language in terms of developing productive language skills. So, although online oral activities may make use of the same methods that face-to-face activities use, the affordances of online activities may make them at least as effective as, and sometimes more practical than, face-to-face activities by increasing the quantity and quality of oral language practice.

Communication can be categorized as either synchronous, having little or no lag time, or asynchronous, having a long lag time, based on Graham’s (2006) description of distance learning environments (see Table 1). Although asynchronous and synchronous communication are similar in some ways, asynchronous communication provides opportunities that synchronous communication (or even classroom speaking activities) does not. First, synchronous communication is more conducive to interpersonal speaking. Ziegler (2016), in her synthesis of synchronous computer-mediated communication (SCMC) use, situated SCMC within the interaction hypothesis, arguing that it provides opportunities for interaction and negotiation of meaning. Asynchronous oral communication, on the other hand, can be considered a type of presentational speaking, a necessary skill in many occupations—see the American Council on the Teaching of Foreign Language’s (2012) description of modes of communication for more information. However, it could be argued that even synchronous conversations consist,

to a degree, of a series of mini-presentations. Whereas Kitade (2000) rightly argued that interlocutors need interaction skills and pragmatic competence when responding to one another in synchronous conversations, they sometimes do so by providing complete, continuous responses or by sharing anecdotes.

Asynchronous	Synchronous
Prepared	Spontaneous
Targets presentational speaking	Targets interpersonal speaking
Disposed to formal evaluation	Disposed to impromptu, informal evaluation
Revisionary	Single occurrence

Second, asynchronous communication more naturally promotes planning before the speech act whereas synchronous communication tends to be more spontaneous. Crookes (1989) discussed the value of pre-task planning to improve non-spontaneous language output. In his study, 40 Japanese learners of English participated in two oral explanation tasks. Group 1 (n=20) was given no preparation and planning time before participating in the task. Group 2 (n=20) was given 10 minutes of preparation and planning before the tasks. Crookes found that learners who planned their output generally produced a greater variety of lexis, more complex language, and more detailed descriptions.

Third, asynchronous communication more naturally allows learners to watch or listen to their own performance and conduct self-evaluation. Instructors and learners in many domains have used video recordings of learner behavior to increase self-awareness and determine what skills they need to focus on. Examples can be found in sports (Hastie, Brock, Mowling, & Eiler, 2012) and medicine (Jamshidi, LeMasters, Eisenberg, Duh & Curet, 2009). In Jamshidi et al.'s (2009) study involving junior surgeons practicing laparoscopic suturing skills, learners benefited from reviewing video recordings of their practice attempts. The learners grew in terms of both self-awareness and skill in part because video recording "provides a matrix of information identical to what was available during the operation itself" (p. 625). This is particularly important in language learning, where the learner's memory is taxed while trying to create a message to the point that they may not be wholly aware of the actual language they are producing. Video provides them with the opportunity to hear exactly what they said. In fact, Jamshidi et al. (2009) argued that this type of video review can not only be used for post-performance assessment but also in pre-performance planning (p. 625).

Fourth, because of its recorded nature, asynchronous communication enables learners to revise and rerecord their performance so that they can publish their best version. Learners have long had the opportunity to improve their composition writing by creating several drafts before submitting a final version. Although, learners can also practice oral presentations before a live audience (e.g., a classmate) or in front of a mirror prior to their final performance, this asynchronous multimedia-based communication (AMOC) provides another outlet for this kind of practice that can be done in the learner's own time. Another benefit that live practice does not afford, however, is that AMOC allows the learner to select the best video or audio draft to submit, rather than having to submit the final performance. Additionally, in some draft-writing processes, learners are even asked to focus on revising a specific element of their writing (e.g., spelling or paragraph structure). Castañeda and Rodríguez-González (2011) incorporated this kind of process in their study of nine university-level learners of Spanish and found that learners increased in terms of speaking, analytic, and evaluation skills.

Although AMOC is generally better suited to promoting self-awareness, revision, and presentational speaking skills, synchronous communication seems to be the more popular of the two in blended language learning environments. It may be easy to think that synchronous communication is better for improving learner speaking proficiency,

given its shorter lag time and better simulation of face-to-face conversations. Because of this, we risk falling into the trap of relegating AMOC to the status of technologies we only use if we do not have bandwidth and hardware that supports synchronous conversation. Yet, given that AMOC provides different affordances than what synchronous communication offers, asynchronous communication can serve different purposes than synchronous communication.

However, even though AMOC *can* provide learners with opportunities to develop their linguistic self-awareness and improve their speaking skills, there is no guarantee that learners *will* make these gains by participating in oral asynchronous activities. The purpose of this literature review, then, is to explore how AMOC has been used to improve speaking skills. Additionally, we examine the methodologies that previous research has used to measure improvements in speaking skills. Thus, in this study we will address the following research questions:

Question 1: What language traits are being promoted with AMOC?

Question 2: What are the challenges to effective use of AMOC?

Question 3: What methods and activities have been used in conjunction with AMOC?

Question 4: What methodologies are commonly used to measure and analyze language gains from using asynchronous multimedia-based oral communication to improve learner speaking skills?

2. Methodology

Literature was located using Academic Search Premier, ERIC, JSTOR, and Scopus. The following combinations of search terms were used: *asynchronous video + language*, *asynchronous CMC + language*, *asynchronous + speaking + language*, *video-mediated communication + language*, *vlog + language*, *Wimba + language*, *oral CMC*, *video drafts + language*, and *blended learning + video + language*. Literature was limited to that published before early 2016.

2.1. Inclusion / exclusion criteria

The following criteria were used to determine which studies to include in this analysis. They are relevance, outlet type, and analysis methods (see Table 2).

Table 2. Summary of inclusion/exclusion criteria	
Criterion	Definition
Relevance	University level learner-created oral asynchronous audio or video productions; research focuses on language gains
Outlet type	Peer reviewed journal articles
Analysis methods	Qualitative and quantitative methods

2.1.1. Relevance

We used the following criteria to determine if studies were sufficiently relevant to this discussion:

- Studies must address asynchronous audio or video communication.
- Videos must be learner created.
- Studies must discuss how learners improved language skills by producing videos.
- Studies must discuss university level class implementation in order to maintain comparability between studies.

2.1.2. Outlet type

Only peer-reviewed journal articles were included in this review. Book chapters and conference proceedings were not included. Conference proceedings, although useful, were not included in order to maintain a higher standard for inclusion in this literature review.

2.1.3. Research type

Only articles including qualitative and quantitative studies were included. This criterion is particularly relevant for research question 1 where both empirical and qualitative information clarified how well learning is taking place. For instance, in Kormos and Dénes' (2004) study, speaking fluency was described in terms of specific, empirical measurements, which enables us to compare fluency across studies. On the other hand, Castañeda and Rodríguez-González (2011) shared learner feedback from self-evaluations after participating in an asynchronous video intervention. While this qualitative data did not provide a clear means of comparing learning effectiveness as did Kormos and Dénes' (2004) study, it did provide insights into the learners' experiences, and it provided other information that might not have been solicited or considered in an empirical study. For instance, one learner discussed the concept of anxiety in their responses (2004), which is an important aspect of the use of asynchronous video communication but would not necessarily be considered in a comparison of fluency gains. Theory and design articles were not included unless they also included either a qualitative or quantitative study showing the effect of their theory or design in practice.

2.1.4. Examples of inclusion/exclusion

Table 3 displays examples of articles found during the literature search along with an indication of whether the example article met a given criterion ("X") or did not meet the criterion ("—"). This is meant to give an explanation of our decision process in choosing which articles to include for review. Of the examples shown in Table 3, only Hirotani and Lyddon (2013) met all three criteria and was, therefore, the only one included in this literature review. Tiraboschi and Iovino (2009) presented activities and a related technology but did not focus on the learning effects of implementing the activities and technology or present any data. Hirotani's (2009) article focused on text-based CMC rather than audio or video CMC. Ono, Onishi, Ishihara, and Yamashiro (2015) presented a paper that was published in the conference proceedings, which did not meet the requirement of being a peer-reviewed journal article. Lamy and Goodfellow (1999) focused on text-based CMC, but also focused on language used during ACMC tasks, rather than language gained from using the tasks.

Example/Non-example	Relevance	Outlet type	Analysis methods	Reason for exclusion
Tiraboschi & Iovino (2009)	—	X	—	No data/design showcase
Hirotani (2009)	—	X	X	Text-based CMC
Hirotani & Lyddon (2013)	X	X	X	NA
Ono, Onishi, Ishihara, & Yamashiro. (2015)	X	—	X	Conference proceeding in book
Lamy & Goodfellow (1999)	—	X	X	Text-based CMC; does not focus on language gains

2.2. Search results

Using the aforementioned search terms and inclusion/exclusion criteria, 22 articles were located (see citations for these articles in the Appendix).

3. Using AMOC in language learning

From this pool of articles, we identified several factors that affect the effectiveness of AMOC activities in language learning contexts. This section begins with a description of the linguistic traits that AMOC activities have been used to improve, then moves to a discussion of challenges inherent in using AMOC, and then concludes with a discussion of the effectiveness of various methods of using AMOC to improve the linguistic traits that will be described.

3.1. Using AMOC to develop specific language traits

In this section, we address the question of *what language traits are being promoted with AMOC*. We will focus on accuracy, fluency, and pronunciation. Although AMOC is used to help learners develop several different linguistic traits, we found that these particular traits need to be treated with more rigor.

3.1.1. Accuracy

By using AMOC, learners are able to increase the accuracy of their speech. In a study on the effects of using AMOC in an ESL writing course, Engin (2014) interviewed participants and analyzed questionnaires, finding that students believed their linguistic accuracy increased as a result of creating their videos. Learners were expected to create English writing explanations (tutorials) for other students in their class in video format. Because of the responsibility of teaching placed upon them and peer dependence on their creating a clear, effective explanation, learners felt compelled to produce linguistically accurate explanations and reduce the number of mistakes in their performance. Engin cited one learner's interview response that the video activity helped their accuracy: "It is a good thing to worry about our English because we improve our English" (2014: 19). Unfortunately, it is not clear in what ways learner speech increased in accuracy nor the learners' basis for determining whether they increased in accuracy or not. Although Engin's findings suggest that AMOC can be used to improve accuracy, additional data and analysis procedures would provide a more rigorous, reliable and trustworthy basis for determining that learner speech became more accurate through producing these videos.

3.1.2. Fluency

Learners using AMOC are also able to develop fluency. In his study of Japanese EFL students, Gromik (2012) found that learners increased their speech rate by 37% over the course of a 13-week video production intervention, comparing average speech production of the first and final weeks. Although the average speech rate of the first week was significantly lower than all subsequent weeks, suggesting that some of the learners' improvement may be attributed to familiarization with the task and the technology, Gromik demonstrated a general increase in speech rate attributable to learner production of asynchronous videos.

Despite the generally positive findings of Gromik's (2012) study, his study leaves us with several questions. For instance, Gromik only considered the speech rate of short videos, where the task limited learners to 30-second video clips. It is unclear whether the learners in this study could sustain this speech rate. It is also unclear whether producing longer videos would offer the same advantage in helping learners develop a higher peak speech rate or a higher consistent speech rate. Gromik also considered only two closely related aspects of fluency: number of words produced and speech rate, or number of words produced per second.

While Gromik's (2012) inclusion of two fluency measures is valuable, it does not represent the wide array of fluency measures available to researchers. In their study on the relationship between proficiency and fluency, Baker-Smemoe, Dewey, Bown, and Martinsen (2014) presented three major categories of speech fluency, each characterized by several different aspects, based on Segalowitz's (2010) work on fluency. These categories are cognitive fluency, perceived fluency, and utterance

fluency. Cognitive fluency refers to the ease with which a speaker is able to create and produce speech; perceived fluency refers to native speaker judgments of how easily the learner produces speech; and utterance fluency refers to measurable aspects of learner speech, including speech rate, hesitations and pausing.

Although Gromik's (2012) study demonstrated the potential value of using AMOC to improve learner fluency, more evidence is needed in order to generalize his findings. Further research should consider the various categories of fluency and the effect of AMOC on fluency in longer videos.

3.1.3. Pronunciation

AMOC has also been shown to help learners develop their pronunciation. In a study involving 39 students of French, Lepore (2014) linked AMOC participation to the learners' perceptions of improvement in their pronunciation. Learners in this study used VoiceThread to produce three audio recordings in response to instructor-created prompts and then commented on one another's recordings. After submitting their recordings, learners completed self-assessments, rating their pronunciation during the recordings.

As with Engin's (2014) findings on increased accuracy, relying solely on the perceptions of untrained learners in Lepore's (2014) study renders the validity of the findings questionable. Although Lepore's self-assessment form provides multiple questions to help the learners think about their pronunciation development (e.g. pronunciation compared to peers' pronunciation, pronunciation improvements as a result of using VoiceThread, and accuracy of specific vowel and consonants in French), it neither provides clear guidance in rating their pronunciation nor provides guidance on what should be rated. In this case, a rubric identifying front rounded vowels, front unrounded vowels, back vowels, and difficult French consonants (e.g. /ʁ/) along with a rating scale, a series of descriptions of performance (e.g. native-like, somewhat native-like), or a series of characteristics (e.g. vowel was not rounded but was at correct height) might guide learners to more accurately and reliably assess their own pronunciation, as well as guide them to improving their pronunciation.

3.1.4. Conclusions about these traits

AMOC has been used to promote language gains in terms of accuracy, fluency, and pronunciation. However, it is not clear what aspects of accuracy were improved through AMOC. For instance, it may be that oral ACMC activities are conducive to lexical accuracy but not syntactic accuracy, or the converse. Fluency seems more clearly affected by AMOC activities, as studies have used more clear and varied measurements to determine fluency gains. Finally, although AMOC was shown to promote pronunciation gains, the evidence supporting this notion is insufficient. This may be remedied through the use of more rigorously developed self-rating systems, through native-speaker raters, or through acoustic measurements, such as comparing learner consonant production with native-speaker production using PRAAT, a popular phonetic analysis program. In summary, AMOC has been shown to have the potential to promote language gains in various linguistic aspects, but additional studies and more rigorous research methods are needed to confirm this.

3.2. *Methods and challenges in using AMOC*

Although AMOC has been shown to be a promising medium for helping learners increase their fluency, accuracy, and pronunciation, the mere inclusion of AMOC in a learning environment does not guarantee these increases. The question remains, then, of how to effectively incorporate AMOC into a course curriculum and how to deal with the challenges that inevitably arise. In this section, we address research question 2 by discussing technological challenges that have arisen in previous studies, and address research question 3 by discussing methods and activities that have contributed to the effective use of AMOC in language learning. The methods and activities discussed are training activities, preparatory activities, project-based learning, and self-evaluation combined with revision.

3.2.1. Technological challenges and training

Although many factors affect the quantity and quality of language learning experiences, whether in a classroom or online, technological challenges in particular affect the learning experience during AMOC activities. A variety of technological challenges exist. Poor internet connection is a common challenge that can be experienced in any location. In their study on Malaysian learners using both audio and video recordings, Bakar, Latiff, and Hamat (2013) reported that even learners at a university experienced connectivity problems, affecting their access to the AMOC activities and thereby their level of participation. Hung's (2012) learners in Taiwan also experienced poor internet.

In addition to internet problems, learners may experience hardware deficiencies and malfunctions. Learners in Bakar, Latiff, and Hamat's (2013) study experienced hardware malfunctions that made it impossible to record their voices. Gleason and Suvorov (2012) stated that their learners also had trouble saving and editing their recordings. In Gromik's (2012) study, some learners were unable to upload video files because they were too large. As these video recordings were 30 seconds or shorter, it seems likely that either some learners were unaware of how to select different codecs and file containers for exporting their video or that the recording software they used did not allow them the option to select different codecs or containers. Hung (2012) confirmed this challenge by stating that his learners had difficulties in converting video files into different formats. This was further complicated by the fact that the vlog (video web log) system used in his study only supported a limited set of file formats. Shih (2010) clarified the problem of file format and file size, adding that internet speed is an important and related factor. Thus, with higher internet speeds, file size may not always be a problem, but with lower internet speeds it will be.

Regarding the problem of access to video recording equipment and editing software, Fukushima (2002) argued that in 2002 the cost of equipment and software licenses was, in fact, not an inhibiting factor for implementing video projects in a language class. By 2018, the affordability and availability of basic editing software and recording equipment has likely increased, leading to better access. This is particularly true when one considers that many university students in the United States own a mobile phone capable of recording high definition videos and performing basic video editing tasks, allowing them to record and edit at any time and in any place. Advanced editing functionality is not necessary for most AMOC tasks, which only require the learner to record a simple video, review it, and then record an additional take rather than splice video segments.

However, because not all learners have mobile phones, or their phones cannot record or edit, it is important to provide other means of recording and editing video files. One way to make recording equipment and editing software available to learners is through university media labs. Some universities offer multimedia labs that loan recording equipment and provide computer stations with editing software. Some even go so far as to offer training in the use of the equipment and software. One drawback to these labs, however, is that they may not provide a suitable environment for recording. As Lepore (2014) stated, a lab setting might lead to some learners reducing their recording quality by speaking softly so as not to disturb other lab users. Background noise might also interfere with recording quality. Despite these drawbacks, labs offer a possible solution to hardware and software challenges, and both learners and instructors are frequently unaware of their existence at their university.

Compounding the technological challenges, many learners do not have sufficient experience using the hardware or software needed to participate in AMOC. Responding to this lack of experience, Bakar, Latiff, and Hamat (2013: 232) stated that their learners would benefit from technical training "so that they are familiar with the online devices and would feel less awkward when utilizing the features of the online tools." One example of this kind of training took place in Abuseileek and Qatawneh's (2013) study where learners were provided with basic instruction in using the AMOC software. Similarly, learners in Fukushima's (2002) study were trained in video and audio editing.

In 2011, Castañeda & Rodríguez-González conducted a study on the effects of self-evaluation and iterative video speech revisions on learners' linguistic self-awareness and

speaking skills. In this study, nine intermediate level Spanish language learners participated in a training activity in which they submitted trial videos prior to participating in the intervention. They created a trial video, following the same procedures they would use to create the videos for the intervention. While the researchers did not mention any specific instruction in how to use the hardware or software, learners nevertheless gained experience in the recording and uploading processes that were required of them in the intervention.

The researchers (Castañeda & Rodríguez-González, 2011) analyzed the learners' self-evaluation forms to determine if learners felt they had made improvement. In their study, Castañeda & Rodríguez-González did not report any learner dissatisfaction with AMOC caused by technological problems. This may be attributed in part to the carefully organized learning activities—where learners participated in four cycles of video recordings and subsequent self-evaluation prior to final submission—but also in part to the technical training learners received.

On the other hand, some learners in Dona, Stover, and Broughton's (2014) study who attended a software training session at the beginning of the course still reported having technological challenges. The researchers cited low learner tolerance for learning new technologies as one cause for this problem, and unclear tutorials as a second. While it is not expected that any training activity would solve all technological challenges, a clear description of the training provided would help in discovering how the training could be clearer and how to adapt the training to learners with low tolerance for new technologies.

In Goulah's (2007) ethnographic case study of eight Japanese language learners, learners were not given any formal training on how to use the recording hardware or editing software. Rather, students with prior experience in recording and editing (whether they gained their experience prior to the course or during the first cycle of the intervention activity) became the experts in the second cycle and assisted other learners at that point. In this case, training was done informally by peers, rather than as a formal instructional session by the instructor or researcher. The value in this approach is that learners may, in fact, learn more from someone with a similar status and may learn more because they are receiving instruction while working with the hardware or software. The danger is that instructors cannot guarantee they will have learners with prior experience, and that it may take learners a much longer time to familiarize themselves with the hardware and software before being able to train their peers.

Although it appears training is valuable in alleviating some technological challenges that learners face, there are different ways of providing that training, and it should be carefully designed. Training may be conducted either formally by the instructor or another expert (Dona, Stover, & Broughton, 2014), or by a more knowledgeable peer (Goulah, 2007). Knowing which learners have prior experience with hardware and software is invaluable if peer-to-peer training is to be expected. Training should also be tailored to the particular learners as much as possible. Many learners are eager to work with new technology, but others are wary of it (Dona, Stover, & Broughton, 2014). Finally, in designing AMOC learning activities, designers must consider learner access to recording hardware and software in the first place. Some may be able to use a mobile phone or personal computer, but others may need access to a lab where they can make their recordings. Yet regardless of the exact nature of the training, training should be provided as many learners lack the skills and equipment necessary to make their recordings, and addressing these deficiencies will help learners to focus on their languaging and not on the technological aspects of the activities.

3.2.2. Preparatory activities

One of the factors that increases the effectiveness of AMOC in developing speaking proficiency is the inclusion of a preparatory activity. Crookes (1989) described planning as a type of preparatory activity in his seminal paper involving 40 Japanese learners of English. He cited "consistent, small- to medium-sized effects in favor of the planned condition" (p. 379), as compared with a control group who did not have planning time. Preparatory activities can take a variety of forms. Bakar, Latiff, and Hamat (2013)

described a simple preparatory activity in which learners were given “time to construct and develop their ideas or thoughts” (p. 232) prior to making their audio and video recordings. This preparation enabled the learners to produce more complex ideas. In order to create their video tutorials, Engin’s (2014) learners conducted their own research on their tutorial subjects, finding, evaluating, selecting, and finally summarizing their sources. This task made the learners responsible for their learning and pushed them to spend time becoming very familiar with it, resulting in students both becoming experts on their topic and developing speaking proficiency.

Goulah (2007) outlined a more complex preparatory activity. Prior to recording their videos, learners in Goulah’s study watched videos related to their video topic and then created a storyboard for their video. The storyboard process involved drafting, presenting, negotiating, and finally settling on ideas as a group. Essentially, learners moved from input, to output, and finally to revision of their output, resulting in exposure to authentic language and more time on task. This kind of preparatory activity takes the focus off languaging, as Knouzi, Swain, Lapkin, & Brooks (2010) use the term, for the sake of language and encourages learners to focus on task completion. Learners were able to experience a real need for language and a purposeful interaction in the target language.

3.2.3. Project-based learning

Incorporating AMOC tasks through project-based learning (PBL) can be an effective method of developing learner speaking skills. PBL does this by creating an authentic need to use the target language and by encouraging learners to use a variety of their target language skills and knowledge. In Goulah’s (2007) study involving eight intermediate learners of Japanese, learners followed a sequence of project-related activities in which they created commercials responding to challenging political and environmental questions. Their project participation resulted in both an increase of content knowledge and language gains.

Fukushima (2002: 353) conducted a study on the effects of PBL in which seven learners collaborated to produce a video promoting Japanese language learning. He described their participation as “self-directed,” highlighting that learners assigned their own tasks, set their own schedule, wrote their own scripts, and evaluated and revised their own performance. The result was that learners produced an authentic linguistic artefact that demonstrated and developed some of their language skills but did not encourage the level of linguistic output and development that the researcher had hoped for. Although language use was considered and reported on, Fukushima focused more attention on motivation and the development of technical skills than on proficiency and performance. A more thorough analysis of the learners’ performance in terms of linguistic dimensions, such as accuracy, fluency, and pronunciation, would allow for comparisons with similar learners and allow for a long-term study analyzing the learners’ linguistic development.

Although neither Goulah’s (2007) nor Fukushima’s (2002) studies suggest PBL as an *efficient* means of bringing about language gains, they both demonstrated that PBL has the potential of creating authentic needs for language learning by motivating learners and giving them opportunities to express themselves. Further studies building on Goulah’s (2007) and Fukushima’s (2002) work should demonstrate ways in which we can efficiently use project-based oral ACMC to create authentic linguistic needs, motivate learners, and bring about significant language gains.

3.2.4. Self-evaluation and revision

In addition to other methods and techniques of incorporating AMOC into learning environments, researchers have found that self-evaluation helps learners achieve language gains. Due to the recorded nature of asynchronous audio and video, learners are not only able to produce spoken output but can listen to their own performance and discover areas of weakness and areas of strength. For instance, most learners in Hung’s (2011) study of Chinese learners of English (76%) agreed that participating in creating vlogs helped them reflect on their learning. One learner described the value of the AMOC project in helping them to become aware of their weaknesses and in being able to make improvements by stating, “I can redo the clips again and again until they looked [sic] satisfactory” (Hung, 2011: 742). Lepore (2014) indicated that self-

evaluation through AMOC was one of the factors involved in increasing learner willingness to communicate, which itself leads to increased quantity of practice. Dixon and Hondo (2014) reported positive learner impressions of the value of AMOC in making them more aware of their speech production, enabling them to make corrections.

In 2011, Castañeda and Rodríguez-González conducted a study in which nine university-level learners of Spanish produced videos of themselves responding to instructor-generated prompts. Learners in this study responded to a prompt by recording an initial video draft, and conducting an evaluation of their draft. They then recorded a second draft and conducted a second self-evaluation. Learners followed this same 2-draft and 2-self-evaluation process, responding to an altered version of the first prompt, although the drafts were labeled as third and fourth drafts. For the self-evaluation, learners watched their recordings, noting mistakes and then recording an improved version.

Learners in Castañeda and Rodríguez-González's (2011: 491) study reported an increase in learner awareness of weaknesses as well as improvements in their grammatical accuracy, pronunciation accuracy, and fluency. Demonstrating increased awareness, one learner stated, "I also noticed my adjective endings weren't correct." Another learner commented on the effect of the self-evaluation and revision cycles, "as we do more recordings, the pauses are becoming less frequent." Castañeda and Rodríguez-González attributed these gains at least in part to the self-evaluation and revision activities.

Of course, incorporating self-evaluation using AMOC does not automatically lead to language gains. Gleason and Suvorov (2012) found that learners were only partially in agreement ($m=3.78$ based on a 5-point scale) that their language skills increased after using AMOC and conducting a self-evaluation. In fact, some learners' perceptions of the value of the intervention actually decreased after participating. In their study, learners recorded three presentations each to share with their peers. They then watched their recordings later to determine if they had made improvements. There is no mention, however, of asking the learners to evaluate their performance and then make changes to their original recording, or to focus on weak areas in subsequent recordings. It seems that learners did not conduct their self-evaluations until after they had completed all their recordings.

Castañeda and Rodríguez-González's (2011) study demonstrated the potential value of combining AMOC with learner self-evaluation and revision cycles. The self-evaluations informed learners of weaknesses and mistakes that learners addressed in subsequent video drafts. Additionally, learners participated in four cycles of self-evaluation and revision. In contrast, learners in Gleason and Suvorov's (2012) study either did not have or did not take the opportunity to improve their recordings based on their self-evaluations. The result was that many did not feel participation in the AMOC activity led to language gains. Thus, while AMOC can be used to create language gains, a structured approach involving both self-evaluation and revision across multiple cycles is more likely to lead to those gains.

3.2.5. Conclusions regarding AMOC methods and challenges

There are a number of things instructors and designers can do to increase the effectiveness of AMOC activities. First, it is important to investigate the learners' hardware and software needs, provide equipment or a lab environment if necessary, and provide training on the creation and sharing of asynchronous audio and video files. If internet speed is a problem, audio might be a more useful option than video, as audio files tend to be much smaller. Second, preparatory activities will improve learner performance. Preparatory activities range in simplicity from brainstorming ideas before recording to viewing related input and then creating a storyboard. Third, project-based learning in AMOC creates authentic needs for learning and encourages learners to be more self-directed. Finally, cycles of structured self-evaluation followed by revisions may raise learners' linguistic self-awareness and provide them with the opportunity to learn from their heightened awareness.

With those benefits in mind, it is important to note that these methods will not guarantee effective and efficient learning through AMOC. Designers and instructors

must incorporate them appropriately, according to the curriculum and the needs of the particular learners. Furthermore, future research is needed to investigate effective methods of incorporating AMOC into a curriculum and to what degree its successful use can be generalized across university-level language learners.

4. Methodologies for measuring and analyzing language gains in AMOC

In this section, we address research question 4. The authors of the articles considered in this review used several methods to determine whether AMOC activities brought about learner language gains. In terms of data type, they analyzed surveys, journals, and reflections; learner audio and video recordings; interview transcripts; and researcher observation notes. Table 4 displays the frequency of use for each data type. In terms of data analysis type, researchers used qualitative analysis, descriptive measurements, quantitative comparison, expert evaluation, and correlation. Table 5 displays the number of studies that used each data analysis type. Each data type and analysis type used by a given study were counted individually. Thus, if a study incorporated surveys, interviews, and recordings, as in Shih (2012), the frequency for surveys, interviews, and recordings would each be increased by one. In this way, the total count for data types and analysis types equaled more than the total number of studies reviewed. Appendix B displays the data and analysis type(s) considered in each study.

Table 4. Frequency of data types	
Data type	Frequency
Surveys, journals, and reflections	16
Audio & video recordings	12
Interview transcripts	10
Observation notes	2

Table 5. Frequency of data analysis types	
Analysis type	Frequency
Qualitative analysis	16
Descriptive measurements	13
Quantitative comparison	6
Expert evaluation	5
Correlation	3
Unknown / unstated	1

4.1. Data sources

Surveys, journals, and reflections was the most common category of data type for determining whether AMOC activities were effective in promoting language gains. Surveys, journals, and reflections were combined into this single category because they

contained the learners' perceptions of their language gains. Many surveys resembled the journals and reflections in that they provided learners with open-ended questions regarding their learning experience, thus increasing the similarity between survey data and journal and reflection data. For instance, Goulah (2007: 65) used surveys to discover that participants felt they learned vocabulary and grammar, referring to his surveys simply as "open-ended questionnaires." Others, however, used surveys to collect data on learner opinions of AMOC technology and activities. One example is Hung's (2011: 742) survey, which largely focused on learner attitudes based on a five-point scale, "the vlog helped me reflect on my learning in this course," though it contained a question related to learner perceptions of language gains "the vlog helped me organise learning in this course."

Interview data, while the third most common of the four categories, resembled survey, journal, and reflection data, differing only in that interviewers personally elicited learner responses rather than providing them with written questions. Like surveys, interviews focused on learner perceptions of language gains (e.g., Kirkgöz, 2011), as well as attitudes (e.g., Hung, 2011; Yaneske & Oates, 2010). In fact, survey and interview data proved to be similar such that many researchers did not state which themes emerged from survey data and which emerged from interview data.

Audio and video recordings were used as a source of data in roughly one half of the studies considered in this review (n=12). Recordings were either coded for qualitative analysis (n= 6), measured and assigned descriptive statistics (n = 4), or assessed using expert evaluation (n= 4). Three studies used two different analysis types on the recordings (Kormos & Dénes, 2004; Sun, 2012; Sun & Yang, 2015).

4.2. Data analyses

Qualitative analysis was the most common data analysis type found in this study. The term qualitative analysis as used in this study refers to any type of coding and categorizing activities. Conversation analysis and discourse analysis were included in this category.

Descriptive measurement was the second most common analysis type. This term refers to frequency counts, means, and standard deviations. It was frequently used in conjunction with qualitative analysis, as in Shih (2010). In his study, Shih counted the frequency of codes found in learner reflections, and calculated means for survey responses. However, some studies provided empirical descriptions of learner language based on their recordings. For instance, Kormos & Dénes (2004: 154) reported 13 statistics, including speech rate, number of words, and mean length of run.

Quantitative comparison refers to quantitative tests used to compare either survey data or learner performance on recordings. In one of the studies (Gromik, 2012), the researcher used a *t*-test to compare learner opinions of the value of using a mobile phone in AMOC activities. In the other five studies using quantitative comparison, the researchers assessed linguistic performance by analyzing recordings and language performance tests. For example, in a study of Turkish learners of English (Kirkgöz, 2011), the means of pre-tests and post-tests were compared using a *t*-test.

Quantitative analysis was used to study the variety of question types and question strategies used (Abuseileek & Qatawneh, 2013); opinions regarding mobile phone use (Gromik, 2012); "fluency, pronunciation, vocabulary, accuracy and task accomplishment" (Kirkgöz, 2011: 4); fluency (13 different measurements) (Kormos & Dénes, 2004); fluency, pronunciation, complexity, and accuracy (Sun, 2012); and pronunciation and grammar (Tognozzi & Truong, 2009).

Expert evaluation refers to either a researcher or instructor's assessment of the learners' performance. For example, Kirkgöz (2011: 4) created a rating scale to assess learner performance in terms of "fluency, pronunciation, vocabulary, accuracy and task completion," which she later used for quantitative comparison. Similarly, in Kormos and Dénes' (2004) study, three native and non-native speakers rated the learners' performance in the AMOC task.

4.3. Conclusions on methodologies

It is puzzling that a majority of studies in this review focused on learner perceptions of language gains without considering expert evaluations or empirical measurements of learner performance. That is, although survey, journal, and reflection data constituted only a marginally larger category than the use of recordings as data, if it were combined with interview data to create the broader category of learner perceptions, it would contain twice as many instances of data collection (n=25) as the recordings category (n=12). It is worth noting that this is a count of instances that each collection method was encountered, where one article may use both surveys and interviews. In other words, researchers relied more heavily on learner perceptions of speech production than on their recorded speech production when studying AMOC in language learning, including studies focusing on the effect of AMOC on learner language gains.

While learner perceptions of linguistic growth and of activity effectiveness are no doubt important aspects in evaluating AMOC and its associated activities, the use of learner perceptions as the sole means of determining this growth and effectiveness is fraught with validity issues. It is doubtful that learners are the best means of gauging language improvement. First, learners are not experts in the language and therefore frequently do not know when they are saying something correctly or incorrectly. Second, they are not trained in noticing different aspects of their own speech. Finally, they are not trained in reliably rating their linguistic performance.

Learner perceptions may still be of value when combined with other analysis methods. One method is expert evaluation. Native speakers and highly proficient non-native speakers are more familiar with the language and can more accurately determine the quality and accuracy of the learner's performance. Objective measurements, such as words produced per second, will provide even more accurate evidence regarding some aspects of learner performance, such as fluency. Taken together, learner perceptions, expert evaluation, and objective measurements would enable researchers to more accurately evaluate learner language gains from using AMOC.

5. Conclusions

AMOC can be beneficial to learners in promoting language gains. Studies considered in this review investigated its effects on accuracy, fluency, and pronunciation, showing that it can be a useful technology in helping learners develop these aspects of their language. However, the research does not universally show that AMOC leads to language gains. Additional studies on the effectiveness of using AMOC would enable us to determine with greater reliability whether it is a viable means of promoting language gains. Additionally, the scope of studies should extend beyond grammatical accuracy, fluency, and pronunciation to include such linguistic aspects as complexity, lexical accuracy, and lexical variety (to name a few).

However, we did identify several factors that contribute to effective use of AMOC in a language-learning curriculum. In designing AMOC activities, instructors and designers should consider the learners' access to hardware and software as well as their internet speed. Because many learners are not familiar with recording and editing software, learners will benefit from technical training. Learners will also benefit from structured self-evaluation and revision cycles, preparatory activities, and project-based learning.

Current research on the effectiveness of AMOC on speaking performance focuses heavily on learner perceptions of language gains. Although learner perceptions can give us clues about their linguistic self-awareness and their experience as AMOC users, they are not an appropriate data source for inferential studies and not the only factor that should be considered by instructors or programs deciding on whether or how to implement AMOC activities. Triangulating with other data sources (such as recordings of learner speech) and other analysis types (such as expert evaluation and empirical measurements) would allow researchers to make more accurate claims as to the effectiveness of AMOC in promoting foreign language gains. This study shows that there are several studies about the qualitative effects of AMOC but few studies providing empirical evidence for linguistic gains through AMOC. What is lacking is an analysis of whether each study's data and analysis type matches the study's claims and

conclusions. Such an analysis would help us to better evaluate the trustworthiness of the various conclusions about the usefulness and effectiveness of AMOC.

In this review, audio-based and video-based AMOC were studied together. However, it is not clear if video-based AMOC is more or less effective at promoting language gains when compared to audio-based AMOC. It is possible that video may be detrimental for some learners in that it will likely increase anxiety when compared to audio. On the other hand, video provides a higher fidelity experience when communicating with other learners or the instructor. A purposeful comparison would help determine if the use of either purely audio or purely video-based AMOC is generally most effective, or to which situations and learner types each is best suited.

A final note is that while self-evaluations and revisions promote language gains, it is unclear what systems for self-evaluating and revising are most effective. For instance, is one cycle of video drafting sufficient or must learners follow three or four cycles before they become sufficiently aware and make sufficient revisions? Furthermore, to what degree do learners even follow the specified self-evaluation and review processes? That is, we do not know the extent to which learners revise their recordings after self-evaluating.

AMOC remains an intriguing means of promoting spoken language gains but further research is needed to determine what aspects of spoken language it is best suited for developing and how to effectively incorporate it into a curriculum. AMOC does not appear to be, as some may think, inferior to face-to-face or other synchronous forms of communication. Continued popularity of asynchronous social media, such as Twitter, Snapchat, and YouTube, suggests that it is important to study and understand the unique outcomes and situations where each method can be most useful.

References

- American Council on the Teaching of Foreign Languages. (2012). *Performance descriptors for language learners*. <http://www.actfl.org/publications/guidelines-and-manuals/actfl-performance-descriptors-language-learners>
- Abuseileek, A. F., & Qataweh, K. (2013). Effects of synchronous and asynchronous computer-mediated communication (CMC) oral conversations on English language learners' discourse functions. *Computers and Education*, 62, 181–190. doi: 10.1016/j.compedu.2012.10.013
- Bakar, N. A., Latiff, H., & Hamat, A. (2013). Enhancing ESL learners speaking skills through asynchronous online discussion forum. *Asian Social Science*, 9(9), 224–234. doi: 10.5539/ass.v9n9p224
- Baker-Smemoe, W., Dewey, D. P., Bown, J., & Martinsen, R. A. (2014). Does measuring L2 utterance fluency equal measuring overall L2 proficiency? Evidence from five languages. *Foreign Language Annals*, 47(4), 707–728. doi: 10.1111/flan.12110
- Castañeda, M., & Rodríguez-González, E. (2011). L2 speaking self-ability perceptions through multiple video speech drafts. *Hispania*, 94(3), 483–501.
- Clark, R. (1994). Media will never influence learning. *Educational Technology Research and Development*, 42(2), 21–29. doi: 10.1152/advan.00094.2010
- Clifford, R. (2002). Achievement, performance, and proficiency testing. Paper presented at the Berkeley Language Center Colloquium on the Oral Proficiency Interview, University of California at Berkeley.
- Crookes, G. (1989). Planning and interlanguage variation. *Studies in Second Language Acquisition*, 11(4), 367–383.
- Delaney, T. (2012). Quality and quantity of oral participation and English proficiency gains. *Language Teaching Research*, 16(4), 467–482. doi: 10.1177/1362168812455586
- Dixon, E. M., & Hondo, J. (2014). Re-purposing an OER for the online language course: A case study of Deutsch Interaktiv by the Deutsche Welle. *Computer Assisted Language Learning*, 27(2), 109–121. doi: 10.1080/09588221.2013.818559

- Dona, E., Stover, S., & Broughton, N. (2014). Modern languages and distance education: Thirteen days in the cloud. *Turkish Online Journal of Distance Education*, 15(3), 155–170.
- Engin, M. (2014). Extending the flipped classroom model: Developing second language writing skills through student-created digital videos. *Journal of the Scholarship of Teaching and Learning*, 14(5), 12–26. doi: 10.14434/josotlv14i5.12829
- Fukushima, T. (2002). Promotional video production in a foreign language course. *Foreign Language Annals*, 35(3), 349–355.
- Gleason, J. & Suvorov, R. (2012). Learner perceptions of asynchronous oral computer-mediated communication: Proficiency and second language selves. *Canadian Journal of Applied Linguistics*, 15(1), 100–121.
- Goulah, J. (2007). Village voices, global visions: Digital video as a transformative foreign language learning tool. *Foreign Language Annals*, 40(1), 62–78. doi: 10.1111/j.1944-9720.2007.tb02854.x
- Gromik, N. A. (2012). Computers & education cell phone video recording feature as a language learning tool: A case study. *Computers & Education*, 58(1), 223–230. doi: 10.1016/j.compedu.2011.06.013
- Graham, C. (2006). Blended learning systems: Definition, current trends, and future directions. In Bonk, C. & Graham, C. (eds.), *Handbook of blended learning: Global perspectives, local designs* (pp. 3–21). San Francisco: Pfeiffer. doi: 10.2307/4022859
- Hastie, P., Brock, S., Mowling, C. & Eiler, K. (2012). Third grade students' self-assessment of basketball dribbling tasks. *Journal of Physical Education and Sport*, 12(4), 427–430. doi: 10.7752/jpes.2012.04063
- Hirotsani, M. (2009). Synchronous versus asynchronous CMC and transfer to Japanese oral performance. *Calico Journal*, 26(2), 413–438. doi: 10.1016/j.cpen.2012.02.001
- Hirotsani, M. & Lyddon, P. A. (2013). The development of L2 Japanese self-introductions in an asynchronous computer-mediated language exchange. *Foreign Language Annals*, 46(3), 469–490. doi: 10.1111/flan.12044
- Hung, S. T. (2011). Pedagogical applications of Vlogs: An investigation into ESP learners' perceptions. *British Journal of Educational Technology*, 42(5), 736–746. doi: 10.1111/j.1467-8535.2010.01086.x
- Jamshidi, R., LaMasters, T., Eisenberg, D., Duh, Q. Y. & Curet, M. (2009). Video self-assessment augments development of videoscopic suturing skill. *Journal of the American College of Surgeons*, 209(5), 622–625. doi: 10.1016/j.jamcollsurg.2009.07.024
- Karweit, N. (1984). Time on task reconsidered: Synthesis of research on time and learning. *Educational Leadership*, 41(8), 32–35.
- Kirkgöz, Y. (2011). A blended learning study on implementing video recorded speaking tasks in task-based classroom instruction. *Turkish Online Journal of Educational Technology*, 10(4), 1–13.
- Kitade, K. (2000). L2 learners' discourse and SLA theories in CMC: Collaborative interaction in internet chat. *Computer Assisted Language Learning*, 13(2), 143–166. doi: 10.1076/0958-8221(200004)13
- Kormos, J. & Dénes, M. (2004). Exploring measures and perceptions of fluency in the speech of second language learners. *System*, 32(2), 145–164. doi: 10.1016/j.system.2004.01.001
- Lamy, M.-N. & Goodfellow, R. (1999). "Reflective conversation" in the virtual classroom. *Language Learning & Technology*, 2(2), 43–61.
- Lepore, C. E. (2014). Influencing students' pronunciation and willingness to communicate through interpersonal audio discussions. *Dimension*, 73–96.

- Lin, H. (2015). Computer-mediated communication (CMC) in L2 oral proficiency development: A meta-analysis. *ReCALL*, 27(3), 261–287. doi: [10.1017/S095834401400041X](https://doi.org/10.1017/S095834401400041X)
- McIntosh, S., Braul, B. & Chao, T. (2003). A case study in asynchronous voice conferencing for language instruction. *Educational Media International*, 40(1), 63–73. doi: [10.1080/0952398032000092125](https://doi.org/10.1080/0952398032000092125)
- Ono, Y., Onishi A., Ishihara M. & Yamashiro M. (2015). Voice-based computer mediated communication for individual practice to increase speaking proficiency: Construction and pilot study. In Zaphiris P. & Ioannou A. (eds.), *Learning and collaboration technologies. LCT 2015. Lecture Notes in Computer Science*, 9192. New York: Springer.
- Pop, A., Tomuletiu, E. A. & David, D. (2011). EFL speaking communication with asynchronous voice tools for adult students. *Procedia - Social and Behavioral Sciences*, 15, 1199–1203. doi: [10.1016/j.sbspro.2011.03.262](https://doi.org/10.1016/j.sbspro.2011.03.262)
- Sauro, S. & Smith, B. (2010). Investigating L2 performance in text chat. *Applied Linguistics*, 31(4), 554–577.
- Segalowitz, N. (2010). *Cognitive bases of second language fluency*. New York: Routledge.
- Shih, R. (2010). Blended learning using video-based blogs: Public speaking for English as a second language students. *Australasian Journal of Educational Technology*, 26(6), 883–897.
- Sun, Y. C. (2012). Examining the effectiveness of extensive speaking practice via voice blogs in a foreign language learning context. *CALICO Journal*, 29(3), 494–506.
- Sun, Y.C. & Yang, F.Y. (2015). I help, therefore, I learn: Service learning on Web 2.0 in an EFL speaking class. *Computer Assisted Language Learning*, 28(3), 202–219. doi: [10.1080/09588221.2013.818555](https://doi.org/10.1080/09588221.2013.818555)
- Tiraboschi, T. & Iovino, D. (2009). Learning a foreign language through the media. *Journal of E-Learning and Knowledge Society*, 5(3), 133–137.
- Tognozzi, E. & Truong, H. (2009). Proficiency and assessment using WIMBA voice technology. *Italica*, 86(1), 1–23.
- Yaneske, E. & Oates, B. (2010). Using voice boards: Pedagogical design, technological implementation, evaluation and reflections. *Australasian Journal of Educational Technology*, 26(8), 233–250. doi: [10.3402/rlt.v18i3.10767](https://doi.org/10.3402/rlt.v18i3.10767)
- Ziegler, N. (2013). *Synchronous computer-mediated communication and interaction: A research synthesis and meta-analysis* (Doctoral dissertation). Washington, DC.

Appendix A

Articles Reviewed in this Study

- Abuseileek, A. F., & Qatawneh, K. (2013). Effects of synchronous and asynchronous computer-mediated communication (CMC) oral conversations on English language learners' discourse functions. *Computers and Education*, 62, 181–190. <http://doi.org/10.1016/j.compedu.2012.10.013>
- Bakar, N. A., Latiff, H., & Hamat, A. (2013). Enhancing ESL learners speaking skills through asynchronous online discussion forum. *Asian Social Science*, 9(9), 224–234. <http://doi.org/10.5539/ass.v9n9p224>
- Castañeda, M., & Rodríguez-González, E. (2011). L2 speaking self-ability perceptions through multiple video speech drafts. *Hispania*, 94(3), 483–501.
- Dixon, E. M., & Hondo, J. (2014). Re-purposing an OER for the online language course: A case study of Deutsch Interaktiv by the Deutsche Welle. *Computer Assisted Language Learning*, 27(2), 109–121. <http://doi.org/10.1080/09588221.2013.818559>
- Dona, E., Stover, S., & Broughton, N. (2014). Modern languages and distance education: Thirteen days in the cloud. *Turkish Online Journal of Distance Education*, 15(3), 155–170.
- Engin, M. (2014). Extending the flipped classroom model: Developing second language writing skills through student-created digital videos. *Journal of the Scholarship of Teaching and Learning*, 14(5), 12–26. <http://doi.org/10.14434/josotlv14i5.12829>
- Fukushima, T. (2002). Promotional video production in a foreign language course. *Foreign Language Annals*, 35(3), 349–355. Retrieved from <http://dx.doi.org/10.1111/j.1944-9720.2002.tb01860.x> <http://search.ebscohost.com/login.aspx?direct=true&db=mz&AN=2002652556&site=ehost-live&scope=site>
- Gleason, J., & Suvorov, R. (2012). Learner perceptions of asynchronous oral computer-mediated communication: Proficiency and second language selves. *Canadian Journal of Applied Linguistics*, 15(1), 100–121. Retrieved from <http://scholar.google.com/scholar?hl=en&btnG=Search&q=intitle:Learner+Perceptions+of+Asynchronous+Oral+Computer-mediated+Communication+:+Proficiency+and+Second+Language+Selves#0>
- Goulah, J. (2007). Village voices, global visions: Digital video as a transformative foreign language learning tool. *Foreign Language Annals*, 40(1), 62–78. <http://doi.org/10.1111/j.1944-9720.2007.tb02854.x>
- Gromik, N. A. (2012). Computers & education cell phone video recording feature as a language learning tool: A case study. *Computers & Education*, 58(1), 223–230. <http://doi.org/10.1016/j.compedu.2011.06.013>
- Hirotsani, M., & Lyddon, P. A. (2013). The development of L2 Japanese self-introductions in an asynchronous computer-mediated language exchange. *Foreign Language Annals*, 46(3), 469–490. <http://doi.org/10.1111/flan.12044>
- Hung, S.-T. (2011). Pedagogical applications of Vlogs: An investigation into ESP learners' perceptions. *British Journal of Educational Technology*, 42(5), 736–746. <http://doi.org/10.1111/j.1467-8535.2010.01086.x>
- Kirkgöz, Y. (2011). A blended learning study on implementing video recorded speaking tasks in task-based classroom instruction. *Turkish Online Journal of Educational Technology*, 10(4), 1–13.
- Kormos, J., & Dénes, M. (2004). Exploring measures and perceptions of fluency in the speech of second language learners. *System*, 32(2), 145–164. <http://doi.org/10.1016/j.system.2004.01.001>
- Lepore, C. E. (2014). Influencing students' pronunciation and willingness to communicate through interpersonal audio discussions. *Dimension*, 73–96.

- McIntosh, S., Braul, B., & Chao, T. (2003). A case study in asynchronous voice conferencing for language instruction. *Educational Media International*, 40(1), 63–73. <http://doi.org/10.1080/0952398032000092125>
- Pop, A., Tomuletiu, E. A., & David, D. (2011). EFL speaking communication with asynchronous voice tools for adult students. *Procedia - Social and Behavioral Sciences*, 15, 1199–1203. <http://doi.org/10.1016/j.sbspro.2011.03.262>
- Shih, R. (2010). Blended learning using video-based blogs: Public speaking for English as a second language students. *Australasian Journal of Educational Technology*, 26(6), 883–897.
- Sun, Y.-C. (2012). Examining the effectiveness of extensive speaking practice via voice blogs in a foreign language learning context. *CALICO Journal*, 29(3), 494–506. Retrieved from <http://proxy1.ncu.edu/login?url=http://search.ebscohost.com/login.aspx?direct=true&db=ehh&AN=75496365&site=eds-live>
- Sun, Y.-C., & Yang, F.-Y. (2015). I help, therefore, I learn: Service learning on Web 2.0 in an EFL speaking class. *Computer Assisted Language Learning*, 28(3), 202–219. <http://doi.org/10.1080/09588221.2013.818555>
- Tognozzi, E., & Truong, H. (2009). Proficiency and assessment using WIMBA voice technology. *Italica*, 86(1), 1–23.
- Yaneske, E., & Oates, B. (2010). Using voice boards: Pedagogical design, technological implementation, evaluation and reflections. *Australasian Journal of Educational Technology*, 26(8), 233–250. <http://doi.org/10.3402/rlt.v18i3.10767>

Appendix B
Comparison of articles reviewed

Author(s)	Data Type					Data Analysis Type				
	Surveys, journals, & reflections	Recordings	Interview transcripts	Observation notes	Qualitative analysis	Descriptive measurements	Quantitative comparison	Expert evaluation	Correlation	Unknown/unstated
Abuseileek, A. F., & Qatawneh, K.		✓			✓		✓			
Bakar, N. A., Latiff, H., & Hamat, A.		✓	✓		✓	✓				
Castañeda, M., & Rodríguez-González, E.	✓				✓	✓				
Dixon, E. M., & Hondo, J.	✓				✓					
Dona, E., Stover, S., & Broughton, N.	✓				✓	✓				
Engin, M.	✓		✓		✓					
Fukushima, T.		✓								✓
Gleason, J., & Suvorov, R.	✓		✓		✓	✓				
Goulah, J.	✓	✓		✓	✓					
Gromik, N. A.	✓	✓	✓		✓	✓			✓	
Hirotsani, M., & Lyddon, P. A.	✓	✓			✓	✓				
Hung, S.-T.	✓	✓	✓		✓	✓				
Kirkgöz, Y.	✓	✓	✓		✓		✓	✓		
Kormos, J., & Dénes, M.		✓				✓	✓		✓	
Lepore, C. E.	✓				✓	✓			✓	
McIntosh, S., Braul, B., & Chao, T.	✓			✓	✓	✓				
Pop, A., Tomuletiu, E. A., & David, D.			✓		✓					
Shih, R.	✓	✓	✓		✓	✓		✓		
Sun, Y.-C.	✓	✓			✓	✓		✓		
Sun, Y.-C., & Yang, F.-Y.	✓	✓	✓		✓	✓		✓		
Tognozzi, E., & Truong, H.	✓				✓		✓	✓		
Yaneske, E., & Oates, B.	✓		✓		✓		✓	✓		

Recommended app

Designing and assessing a digital, discipline-specific literacy assessment tool

Paul Graham Kebble

National Institute of Education Nanyang Technological University, Singapore

paul.kebble@nie.edu.sg

Abstract

The C-Test as a tool for assessing language competence has been in existence for nearly 40 years, having been designed by Professors Klein-Braley and Raatz for implementation in German and English. Much research has been conducted over the ensuing years, particularly in regards to reliability and construct validity, for which it is reported to perform reliably and in multiple languages. The author engaged in C-Test research in 1995 focusing on concurrent, predictive and face validity. Through this research, the author developed an appreciation for the C-Test assessment process particularly with the multiple cognitive and linguistic test-taking strategies required. When digital technologies became accessible, versatile and societally integrated, the author believed the C-Test would function well in this environment. This conviction prompted a series of investigations into the development and assessment of a digital C-Test design to be utilised in multiple linguistic settings. This paper describes the protracted design process, concluding with the publication of mobile apps.

Keywords: C-Test, language competence assessment, mobile app design.

1. Introduction

Over the past eight years, I have been involved in the design and implementation and evaluation of a series of digital language competency assessment tools utilising a unique testing format, the C-Test. The C-Test construct (Raatz, 1985a; Raatz & Klein-Braley, 1985; 2002) allows for the design of a reliable and versatile (Raatz, 1987; Rouhani, 2008) subject-specific language competency assessment tool (Alderson, 2002; Grotjahn, 1987; Grotjahn & Stemmer, 2002; Hulstijn, 2010) through the utilisation of subject-related texts. A digital C-Test, I conjectured, could be constructed to provide a linguistic assessment tool relevant to any academic or professional discipline, and highly efficient in implementation. Output of such a test would allow assessors to identify takers' levels of linguistic competency and indicate those who might require linguistic support. For example, a lecturer teaching into a 1st year Electrical Engineering subject could use the online C-Test, utilising texts selected from the 1st year Electrical Engineering program, to identify students who require subject-specific language support. An international finance company recruiting staff would be able to use a C-Test designed to determine linguistic suitability of potential employees. Equally, the C-Test could utilise texts from a year 5 syllabus and could be administered on a regular basis (bi-monthly), providing a teacher with regular assessment of language and literacy levels of her students, and hence an individual's development. My aim therefore was to design a software application which would efficiently and effectively build the test through the uploading of discipline-specific texts, and assess, collate, analyse and distribute results as to pre-specified requirements.

2. The C-Test

The C-Test construct was originally designed by (the late) Professor Christine Klein-Braley and Professor Ulrich Raatz, both of the University of Duisburg, in the 1980s and was initially utilised for assessing both German (Klein-Braley, 1985c) and English language competency. During this period, professors Klein-Braley and Raatz conducted

extensive research into the validity and reliability of the C-Test, along with many other academics and researchers, such as Professor Grotjahn and Professor Coleman, (Klein-Braley & Raatz, 1984) and found the C-Test was a highly reliable assessor of linguistic competence. In 1996, I was very fortunate to meet both Professors at a language testing conference held at the University of Portsmouth (UK) Subsequently, Prof. Raatz has kindly given full permission for the C-Test construct to be used for the basis of my online design.

The C-Test assesses linguistic competence through reduced redundancy (Klein-Braley, 1985d, 1985e, 1997; Oscarson, 1991; Raatz, 1985b, c) and the accurate restoration (Babaii & Fatahi-Majd, 2014) of a text, where interference in communication is achieved through systematic word mutilation. The C-Test is a derivation of a cloze test where every nth word, usually the 7th, is removed, however, the C-Test textual restoration functions at word, not sentence level, and is described as the 'rule of two' (Jafarpur, 1999). The C-Test is built utilising short texts, usually paragraphs taken from authentic sources (Atai & Soleimany, 2009; Khodadady, 2013; Khodadady & Hashemi, 2011; Klein-Braley, 1985a), with paragraphs obtained from either a variety of genres or one specific genre, depending on the assessment requirements (Mochizuki, A., 1994). From the beginning of the second sentence of each paragraph, the second half of every second word is removed (the rule of two), until 25 deletions are achieved, with the remaining text left intact. A mutilated text looks like this:

Given the continuing changes in the application of Information and Communication Technology (ICT) this survey focuses on broadband access and types and penetration rates of electronic commerce (e-commerce; the online sale of goods and services). Despite t __ growing le __ of pub __ debate conce __ the gro __ of e-bus __ and e-com __ within t __ Australian eco __, very lit __ data rela __ to t __ business u __ of th __ facilities h __ been avai __ to pol __ makers o __ private enter __. Data th __ are avai __ have predom __ been gene __ by pri __ surveys wi __ small samples and inconsistent definitions and scope. At best, the results of these surveys are indicative. There is a pressing need for the production of timely and comprehensive e-commerce statistics.

The original text:

Given the continuing changes in the application of Information and Communication Technology (ICT) this survey focuses on broadband access and types and penetration rates of electronic commerce (e-commerce; the online sale of goods and services). Despite the growing level of public debate concerning the growth of e-business and e-commerce within the Australian economy, very little data relating to the business use of these facilities has been available to policy makers or private enterprises. Data that are available have predominantly been generated by private surveys with small samples and inconsistent definitions and scope. At best, the results of these surveys are indicative. There is a pressing need for the production of timely and comprehensive e-commerce statistics.

Micro- and macro-level textual cues, along with anaphoric and cataphoric referencing, are specific linguistic skills required for restoration, the more accurate the restoration, the more proficient the restorer (Klein-Braley, 1996; Wedell, 1987). Micro-level textual cues refer to an individual word's construct both in terms of word composition and form, whereas Macro-level cues require understanding of syntax and function. Anaphoric referencing necessitates the test taker to refer back within the text, and cataphoric referencing, forward. With the first sentence and the remaining text left intact, the test taker is also able to contextualise the extract, providing further clues for its restoration.

Over the past 30 years, extensive research has been conducted into the C-Test procedure, its reliability and validity (Grotjahn, 1986; Khodadady, 2014; Klein-Braley, 1985b), and its function in a variety of languages. A link to a comprehensive bibliography of research is provided by Professor Rudiger Grotjahn (Grotjahn, 2014), at Ruhr-Universität Bochum:

<http://www.C-Test.de/deutsch/index.php?lang=de&content=bibliografie§ion=ctest>

3. Initial involvement with the C-Test

I first became aware of the C-Test through Master's research conducted whilst at the University of Portsmouth, U.K. (1995-6). For my thesis, I was invited by Professors Rastall and Coleman, both at the University of Portsmouth, to assess the face, concurrent and predictive validity of the C-Test, at the time a paper test being used as the university's language assessment tool for international students entering undergraduate courses. Much research had already been conducted into the C-Test's use in higher education (Coleman, 1994a/b; Coleman, Grotjahn, Rüdiger & Ratz, 2002) with its specific use as a placement test being the focus for my research. With access to students' academic records, I was able to determine that the C-Test had no predictive capability. In terms of concurrent validity, I measured the C-Test against a well-validated commercially available test (the Oxford Placement Test) through the testing of 83 students with both. A strong correlation coefficient (Pearson's *r*) was seen (0.8) allowing my research to suggest the C-Test was a valid test of general language proficiency. This research also indicated that the C-Test suffered from poor face validity, through collated and analysed answers to a questionnaire, analysis showed test takers were not convinced the C-Test functioned as its purpose was described.

With the advent of the internet, I conjectured the C-Test design structure could be highly functional in, and would be especially suited to, a digital environment. My initial attempts at providing the test digitally were limited primarily to attempting a build within the Moodle and Blackboard online learning platforms. And the first time I was able to effectively use such a build was in 2011.

4. Developing the digital C-Test

In 2011, I created an Online C-Test within the Blackboard learning platform when required by my Australian university (James Cook University) to assess the academic language proficiency of a large group (93) of Chinese undergraduate I.T. students studying at a Beijing university on a course provided by my university and in English medium. Through an external quality audit by the Australian Universities Quality Agency (AUQA), concern had been raised at the level of English demonstrated by some students, and my university wished to assess English language levels. Although students were required to have reached an IELTS overall 5 score, or its equivalent, to enter the course, concern had been raised that a large proportion of the 1st year cohort were not functioning acceptably in the learning and teaching medium of English. I was required to design and provide an English language competency test that could be administered efficiently, and did so using a C-Test construct in Blackboard, with four texts sourced from the 1st year IT syllabus.



Figure 1: C-Test build in Blackboard.

As a benchmark, the same C-Test was administered to a group of eight English language students in a private language college in Cairns, Australia. The Cairns students had IELTS scores ranging from 5 to 6.5. The range of scores from these students on the

same C-Test were from 43 to 78. The C-Test results were aligned to the IELTS scores resulting in an informed approximation that scores on the C-Test >40 = IELTS 5, >50 = 5.5, >60 = 6, >70 = 6.5. When this alignment was applied to the results of the Chinese students' C-Test results the indication was that students scoring below 40 might be identified as being at risk of failing to achieve an IELTS 6 on the end-of-year IELTS equivalency test. The results indicated that this amounted to 36 (39%) students out of the total cohort.

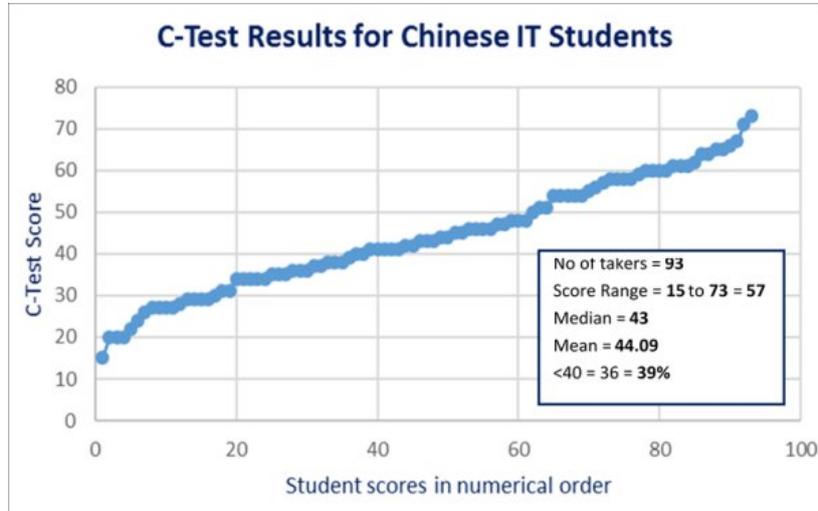


Figure 2: C-Test scores for Chinese IT students.

Although the testing process functioned well, and provided invaluable information, I was dissatisfied with the online test build in Blackboard as the user interaction was not as effective as a paper-based test would have been. Most importantly, with the Blackboard build, takers were unable to complete each word in situ (see diagram 1). Of course, there were other advantages, not using paper being one, and the automatic collation of test results and scores being the other. On returning to my university, I enquired of the I.T. department whether the C-Test could be built in another platform allowing user interaction to match a paper version. I was told it could be done, but to purchase the software and employ a programmer to build the required code would be inordinately expensive.

A move to the University of Tasmania, and meeting D. Heidermann, a senior IT manager, provided expertise in programming, and an interest in the concept. Our initial C-Test build was online and allowed paragraphs of text to be uploaded, multiple tests taken, and results to be collated and distributed. Soon after completing this package, we considered designing a mobile app that would assess language competency and instantly provide the taker with a result and related feedback. Over the ensuing year, we were able to develop and publish three discipline-specific language and literacy assessment apps, initially with Apple, and more recently as an Android app.

5. The 'IELTS Score Predictor' App

My initial design concept was for a mobile app which would provide a predicted overall score for the International English Language Testing System (IELTS). The IELTS examination is a highly reliable and very well validated English language assessment tool, providing a score for each of the macro language skills, reading writing, speaking and listening, and an averaged overall score ranging from 2 to 9, with 0.5 graduations. Undergraduate university courses usually require a 6 or 6.5 overall score, with postgraduate studies requiring 7 or 7.5. My concept was to offer a tool that would provide an overall indicator of an averaged IELTS score, along with a linguistic description of what the achieved score aligned to in terms of functionality in English. I was aware of the limitations of not providing scores for each linguistic skill, but was convinced that receiving an indication of an overall score was highly beneficial.

The app would utilise paragraphs from texts used in past IELTS papers, freely available online. Informed through the previously described use of the C-Test for achieving this

outcome, I worked with D. Heidermann to design and build the software, create the C-Test paragraphs and formulate a rubric to which each score would align.

Originally, the C-Test design used four texts, however, for the IELTS predictor app I decided to use 6. Four texts would be sourced from various sections of IELTS reading tests adapted from past papers; one text would be created from a transcript of an item in the listening section; and the final text would be an imagined dialogue between a test taker and an IELTS interlocutor, achieved through many years of experience teaching and examining IELTS takers. For score alignment, the test was trialled at the Cairns language school, Queensland, Australia, (Cairns College of English) with results employed to create a table of score equivalency, shown in Table 1.

Band Score	9	8.5	8	7.5	7	6.5	6	5.5	5	4.5	4	3.5	3	2.5
% Score	91-100	86-90	80-85	75-79	70-74	64-69	57-63	47-56	37-46	32-36	25-31	20-24	15-19	10-14

Cognisant of my earlier research into face validity, I decided to provide test takers with a concise introduction to the C-Test construct and development, with a link to the bibliography of related research. Clear instructions were also provided to test takers prior to engagement with the C-Test, advising: 1) Read the first sentence carefully and think about the topic of the text, 2) Look at the first mutilated word, in the example above 'i____' and the word before and after. If you can repair the word from these clues, do so, 3) Sometimes you will need to think back to the first sentence, look back in the sentence, and look forward through the sentence, 4) read back through the completed sentences to check if the words sound or appear suitable.

Although the results were graduated at 0.5, the language level descriptors were provided for each integer. If a score was at 0.5, the taker would be required to engage with both higher and lower integer descriptors. Descriptors (Table 2) were written with reference to the British Council's public reference band score descriptors (<http://takeielts.britishcouncil.org/find-out-about-results/ielts-assessment-criteria>).

7 Good user	You have an operational command of the language, though with occasional inaccuracies, inappropriate usage and misunderstandings in some situations. Generally, you handle complex language well and understand detailed reasoning.	You can speak and engage in conversation and discussion at length without noticeable effort or loss of coherence. You may demonstrate language related hesitation at times, or some repetition and/or self-correction, or may ask for clarification when something is not clear. You use a range of connectives and discourse markers with some flexibility. You have little problem in understanding the vast majority of language engaged with, either with other people or through media.	You are able to logically organise information and ideas and there is clear progression throughout. You can use a range of cohesive features* appropriately, although there may be some mistaken use, particularly prepositions. You are able to use and understand a sufficient range of vocabulary to allow some flexibility and precision. You are able to use less common lexical items with some awareness of style and collocation, but you may produce occasional errors in word choice, spelling and/or word formation.
6 Competent user	Generally you have an effective command of the language despite some inaccuracies, inappropriate usage and misunderstandings. You can use and understand fairly complex language, particularly in familiar situations.	You are willing to speak at length, and engage with others, although you may lose coherence at times due to occasional repetition, self-correction or hesitation. You can use a range of connectives and discourse markers, but not always appropriately. You have few problems understanding sympathetic communicators, but	You can arrange information and ideas logically and coherently and there is a clear overall development of ideas and thoughts. You can use a range of common cohesive features* effectively, but cohesion of text within and/or between sentences may be faulty or mechanical. You may not always use referencing clearly or appropriately. You are

		occasionally need to ask for repetition or clarification.	able to use a satisfactory range of vocabulary and you attempt to use less common vocabulary, but with some mistakes. You sometimes make errors in spelling and/or word formation, but they do not cause too many problems. You understand much of general texts, but often need a dictionary or to ask for clarification if reading academic or scientific texts.
--	--	---	---

6. The 'How good is my English?' Apps

These two apps, although fundamentally the same, were designed either to be used by those learning and studying English as a 2nd or other language (E4L2) and those who use English as a first language (E4L1). The E4L1 and E4L2 C-Tests use 8 paragraphs of increasingly complex language taken from multiple sources, ranging from children's literature to academic texts, from newspaper articles to non-fiction prose. Although the C-Test is the same, the output score and descriptors differ. The E4L2 test score aligns with language learning levels widely used by English language learning course books and schools (Table 3), whilst the E4L1 results align with globally recognised education systems and levels (Table 3). Score alignment to descriptors in E4L1 provides takers with a basic language proficiency explanation for their specific level, as shown below (Table 4, whilst E4L2 takers are provided with descriptors more appropriate for language learners (Table 5).

C-Test results for E4L2 takers		C-Test results for E4L1 takers	
Percentage %	Level	Percentage %	Level
0-10	Beginner	0-29	Below year 6
11-29	Elementary	30-40	Year 7-8
30-45	Pre-intermediate	41-55	Year 9-10
46-59	Intermediate	56-64	Year 11-12 & Diploma
60-69	Post-intermediate	65-77	Undergraduate Degree
70-79	Pre-advanced	78-87	Master's Degree
80-89	Advanced	88-100	Doctoral Degree
90-100	Near-native like user		

Table 4. Descriptor for post-intermediate level in E4L2 C-Test feedback

Score of 60-69 - Post-intermediate level:

Generally, you have an effective command of the language despite some inaccuracies, inappropriate usage and misunderstandings. You can use and understand fairly complex language, particularly in familiar situations. You are willing to speak at length, and engage with others, although you may lose coherence at times due to occasional repetition, self-correction or hesitation. You can use a range of connectives and discourse markers, but not always appropriately. You have few problems understanding others, and occasionally need to ask for repetition or clarification. You can arrange information and ideas logically and coherently and there is a clear overall development of ideas and thoughts. You can use a range of common cohesive features* effectively, but cohesion of text within and/or between sentences may be faulty or mechanical. You may not always use referencing clearly or appropriately. You are able to use a satisfactory range of vocabulary and you attempt to use less common vocabulary, but with some mistakes. You sometimes make errors in spelling and/or word formation, but they do not cause too many problems. You understand much of general texts, but often need a dictionary or to ask for clarification if reading academic or scientific texts.

Note: * requires reader to refer to post-script note.

Table 5. Descriptor for undergraduate degree level in E4L1 C-Test feedback

Score of 65-77 - Undergraduate degree level:

You have a good operational command of the language, though with occasional inappropriate usage and misunderstandings in some unfamiliar situations. Generally, you handle complex language well and understand detailed reasoning. You can speak and engage in conversation and discussion at length without noticeable effort or loss of coherence. You use a wide range of connectives and discourse markers with flexibility. You have little problem in understanding the vast majority of language engaged with, either with other people or through media, however, you may find some academic texts more difficult to engage with and will need to read multiple times to have a clear understanding. You are able to logically organise information and ideas and there is clear progression throughout. You can produce quality text that meets most academic and professional requirements.

7. Conclusion

The IELTS predictor was published in February 2017 and by August had had close to 2000 downloads. Being able to track, through Apple statistics, where downloads have occurred is valuable, to date with approximately 80% of downloads in the Asia Pacific region, 9% in the Americas, and 5% in Europe, with China leading by far on a country by country basis. The 'How good is your English' app has had somewhat fewer downloads in its 3 months presence. All three apps will be published in Android form by the end of 2017, and all are free to download. All apps require further research, particularly in terms of score alignments, as only preliminary appraisal has been conducted prior to the apps build and publishing. My programming colleague and I will now concentrate on further refining the initial online C-Test software, and wish also to engage extensively in research around its effectiveness and practicality, particularly within diverse learning, academic and professional settings, and I would be delighted to engage with anyone who might be interested in conducting further research within their specific setting.

Downloads - 26 Feb to 26 Aug - top 10 Territories	
Territory	Units
China	1318
USA	118
India	45
Australia	34
UK	33
Saudi Arabia	11
Thailand	8
Nigeria	6
Philippines	5
France	5

Figure 3. Downloads of The IELTS Score Predictor.

References

- Alderson, J.C. (2002). Testing proficiency and achievement: principles and practice. In James A. Coleman, Rüdiger Grotjahn & Ulrich Raatz (Eds.), *University language testing and the C-test* (pp. 15-30). Bochum: AKS-Verlag.
- Atai, M. R. & Soleimany, M. (2009). On the effect of text authenticity & genre on EFL learners' performance in C-tests. *Pazhuhesh-e Zabanhaye Khareji* 49, 109-123. Available from https://jor.ut.ac.ir/m/article_27676_b6d42cffb07714a23474fe66363c8eee.pdf.
- Babaii, E. & Fatahi-Majd, M. (2014). Failed restorations in the Ctest: Types, sources, and implications for C-test processing. In Rüdiger Grotjahn (Ed.), *Der C-Test: Aktuelle Tendenzen/ The C-Test: Current trends* (pp. 263-276). Frankfurt am Main: Lang.
- Coleman, J. A. (1994a). Degrees of proficiency: Assessing the progress and achievement of university language learners. *French Studies Bulletin*, 50, 11-16.
- Coleman, J. A. (1994b). Profiling the advanced language learner: The C-Test in British further and higher education. In Rüdiger Grotjahn (Ed.), *Der CTest. Theoretische Grundlagen und praktische Anwendungen* (Vol. 2, pp. 217- 237). Bochum: Brockmeyer. Available from <http://www.c-test.de/deutsch/index.php?lang=de§ion=originalia>
- Coleman, J. A., Grotjahn, Rüdiger & Raatz, Ulrich. (Eds.). (2002). *University language testing and the C-test*. Bochum: AKS-Verlag.
- Grotjahn, R. (1986). Test validation and cognitive psychology: some methodological considerations. *Language Testing*, 3(2), 159-185.
- Grotjahn, R. (1987). How to construct and evaluate a C-Test: A discussion of some problems and some statistical analyses. In Rüdiger Grotjahn, Christine Klein-Braley & Douglas K. Stevenson (Eds.), *Taking their measure: The validity and validation of language tests* (pp. 219-253). Bochum: Brockmeyer. Available from <http://www.c-test.de/deutsch/index.php?lang=de§ion=originalia>
- Grotjahn, R. (2014). The C-Test bibliography: version January 2014. In Rüdiger Grotjahn (Ed.), *Der C-Test: Aktuelle Tendenzen/The C-Test: Current trends* (pp. 325-363). Frankfurt am Main: Lang.
- Grotjahn, R. & Stemmer, B. (2002). C-Tests and language processing. In James A. Coleman, Rüdiger Grotjahn & Ulrich Raatz (Eds.), *University language testing and the C-test* (pp. 115-130). Bochum: AKS-Verlag. Available from <http://www.c-test.de/deutsch/index.php?lang=de§ion=originalia>
- Hulstijn, J. H. (2010). Measuring second language proficiency. In Elma Blom & Sharon Unsworth (Eds.), *Experimental methods in language acquisition research* (pp. 185-199). Amsterdam: Benjamins. [C-Test: 191-193]
- Jafarpur, A. (1999). What's magical about the rule-of two for constructing C-tests? *RELC Journal*, 30(2), 86-100.

- Khodadady, E. (2013). Authenticity and sampling in C-Tests: A schema-based and statistical response to Grotjahn's critique. *The International Journal of Language Learning and Applied Linguistics World*, 2(1), 1-17.
- Khodadady, E. (2014). Construct validity of C-Tests: A factorial approach. *Journal of Language Teaching and Research*, 5(6), 1353-1362. Available from <http://www.academypublication.com/issues/past/jltr/vol05/06/19.pdf>
- Khodadady, E. & Hashemi, M. (2011). Validity and C-Tests: The role of text authenticity. *Iranian Journal of Language Testing*, 1(1), 30-41. Available from <http://www.ijlt.ir/portal/files/404-2011-01-01.pdf>
- Klein-Braley, C. (1985a). A cloze-up on the C-test: A study in the construct validation of authentic tests. *Language Testing*, 2(1), 76-104.
- Klein-Braley, C. (1985b). C-Tests and construct validity. In Christine Klein-Braley & Ulrich Raatz (Eds.), *Fremdsprachen und Hochschule 13/14: Thematischer Teil: C-Tests in der Praxis* (pp. 55-65). Bochum: AKS-Verlag.
- Klein-Braley, C. (1985c). C-Tests as placement tests for German university students of English. In Christine Klein-Braley & Ulrich Raatz (Eds.), *Fremdsprachen und Hochschule 13/14: Thematischer Teil: C-Tests in der Praxis* (pp. 96-100). Bochum (Ruhr-Universität): AKS-Verlag
- Klein-Braley, C. (1985d). Reduced redundancy as an approach to language testing. In Christine Klein-Braley & Ulrich Raatz (Eds.), *Fremdsprachen und Hochschule 13/14: Thematischer Teil: C-Tests in der Praxis* (pp. 1-13). Bochum (Ruhr-Universität): AKS-Verlag.
- Klein-Braley, C. (1985e). Tests of reduced redundancy – theory. In Viljo Kohonen & Antti J. Pitkänen (Eds.), *Language testing in school. AFinLA Yearbook 1985* (pp. 33-48). Tampere: AFinLA.
- Klein-Braley, C. (1996). Towards a theory of C-Test processing. In Rüdiger Grotjahn (Ed.), *Der C-Test. Theoretische Grundlagen und praktische Anwendungen* (Vol. 3, pp. 23-94). Bochum: Brockmeyer. Available from <http://www.c-test.de/deutsch/index.php?lang=de§ion=originalia>
- Klein-Braley, C. (1997). C-Tests in the context of reduced redundancy testing: An appraisal. *Language Testing*, 14(1), 47-84.
- Klein-Braley, C. & Raatz, U. (1984). A survey of research on the C-Test. *Language Testing*, 1(2), 134-146.
- Mochizuki, A. (1994). C-Tests: Four kinds of texts, their reliability and validity. *JALT Journal*, 16(1), 41-54.
- Oscarson, M. (1991). Item response theory and reduced redundancy techniques: Some notes on recent developments in language testing. In Kees de Bot, Ralph B. Ginsberg & Claire Kramsch (Eds.), *Foreign language research in cross-cultural perspective* (pp. 95-111). Amsterdam & Philadelphia: Benjamins [C-Test: pp. 106-107].
- Raatz, U. (1985a). Better theory for better tests? *Language Testing*, 2, 60- 75.
- Raatz, U. (1985b). Tests of reduced redundancy – the C-test, a practical example. In Viljo Kohonen & Antti J. Pitkänen (Eds.), *Language testing in school. AFinLA Yearbook 1985*, No. 41 (pp. 49-62). Tampere: Association Finlandaise de Linguistique Appliquée. [long version of Raatz, 1985e]
- Raatz, U. (1985c). Tests of reduced redundancy – the C-Test, a practical example. In Christine Klein-Braley & Ulrich Raatz (Eds.), *Fremdsprachen und Hochschule, 13/14: Thematischer Teil: C-Tests in der Praxis* (pp. 14-19). Bochum: AKS-Verlag
- Raatz, U. (1987). C-Tests: What do they measure, how do they measure it and what can they be used for? In Christine Schwarzer & Bettina Seipp (Eds.), *Trends in European educational research* (pp. 38-44). Braunschweig: Universität

Raatz, U. & Klein-Braley, C. (1985). How to develop a C-Test. In Christine Klein-Braley & Ulrich Raatz (Eds.), *Fremdsprachen und Hochschule, 13/14: Thematischer Teil: C-Tests in der Praxis* (pp. 20-22). Bochum: AKSVerlag.

Raatz, U. & Klein-Braley, C. (2002). Introduction to language testing and to C-Tests. In James A. Coleman, Rüdiger Grotjahn & Ulrich Raatz (Eds.), *University language testing and the C-test* (pp. 75-91). Bochum: AKS-Verlag. Available from <http://www.c-test.de/deutsch/index.php?lang=de§ion=originalia>

Rouhani, M. (2008). Another look at the C-Test: A validation study with Iranian EFL learners. *The Asian EFL Journal*, 10(1), 154-180. Available from [http://www.asian-efl-journal.com/March_2008_EBook.pdf]

Wedell, M. (1987). The C-test - current relevance, text difficulty and student strategies. *Language Testing Update*, 4, 23. [summary of Wedell, 1985].

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 Politècnica 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 Rüschoff (University of Duisburg-Essen, DE) | Dr. Shannon Sauro (Malmö University, SE) | Dr. Mathias Schulze (University of Waterloo, Canada) | Dr. Rafael Seiz-Ortiz (Universidad Politècnica 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 science.
-  Dialnet *plus*
-  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.

If you can provide information on any topic of interest to the CALL community, please send this directly to Prof. Ana Gimeno, email agimeno@upvnet.upv.es

The EUROCALL Review, edited by Prof. Ana Gimeno, is licensed under a [Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License](https://creativecommons.org/licenses/by-nc-nd/4.0/). 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.

