# Table of Contents

Comparing the efficacy of digital flashcards versus paper flashcards to improve receptive and productive L2 vocabulary  
Gilbert Dizon and Daniel Tang .......................................................... 3

A freely available authoring system for browser-based CALL with speech recognition  
Myles O’Brien .................................................................................. 16

Using Facebook to improve L2 German students’ socio-pragmatic skills  
Axel Harting .................................................................................... 26

Testing audiovisual comprehension tasks with questions embedded in videos as subtitles: a pilot multimethod study  
Juan Carlos Casañ Núñez .................................................................. 36

Profiling language learners in hybrid learning contexts: Learners’ perceptions  
Pekka Lintunen, Maarit Mutta and Sanna Pelttari ................................. 61

The perceptions of a situated learning experience mediated by novice teachers’ autonomy  
Paul Booth, Isabelle Guinmard and Elizabeth Lloyd ............................. 76

Editorial Board .................................................................................. 92
Research paper

Comparing the efficacy of digital flashcards versus paper flashcards to improve receptive and productive L2 vocabulary

Gilbert Dizon* and Daniel Tang**
*Himeji Dokkyo University, Japan | Otemae University, Japan
**dtang@otemae.ac.jp

Abstract

Several researchers have compared the efficacy of digital flashcards (DFs) versus paper flashcards (PFs) to improve L2 vocabulary and have concluded that using DFs is more effective (Azabdaftari & Mozaheb, 2012; Başoğlu & Akdemir, 2010; Kiliçkaya & Krajka, 2010). However, these studies did not utilize vocabulary learning strategies (VLSs) as a way to support the vocabulary development of those using PFs. This is significant because DFs often offer a range of features to promote vocabulary development, whereas PFs are much more basic; thus, learners who study via paper materials are at a disadvantage compared with those who use DFs. Given the success that VLSs have had in fostering L2 vocabulary enhancement (e.g., Mizumoto & Takeuchi, 2009), their incorporation could have influenced the previous studies. Therefore, one of the primary aims of this study was to find if there were significant differences in receptive and productive L2 vocabulary improvements between students who used PFs in conjunction with 3 VLSs – dropping, association, and oral rehearsal – and those who used the DF tools Quizlet and Cram. Additionally, the researchers examined the learners’ opinions to see if there was a preference for either study method. A total of 52 EFL students at two Japanese universities participated in the 12-week study. Pre- and post-tests were administered to measure the vocabulary gains in the PF group (n = 26) and the DF group (n = 26). Results from a paired t-test revealed that both groups made significant improvements in receptive and productive vocabulary. However, the difference between the gains was not significant, which contrasts with past comparison studies of DFs and PFs and highlights the importance of VLSs. A 10-item survey with closed and Likert-scale questions was also administered to determine the participants’ opinions towards the study methods. Higher levels of agreement were found in the experimental group, indicating that the students viewed DFs more favorably than PFs.

Keywords: L2 vocabulary, flashcards, computer-assisted language learning, EFL.

1. Introduction

Technological advancements have touched on every aspect of our lives, and vocabulary acquisition is no exception. Increasing affordability and accessibility to the Internet and personal computing devices now means learners have a wide variety of digital flashcard (DF) tools, such as Quizlet and Cram, at their disposal. However, research into the efficacy of DFs versus paper flashcards (PFs) in L2 vocabulary acquisition has not been widespread. Of these studies, most have found DFs to be more effective than PFs (Azabdaftari & Mozaheb, 2012; Başoğlu & Akdemir, 2010; Kiliçkaya & Krajka, 2010); however, these studies did not incorporate vocabulary learning strategies (VLSs), which might have leveled the playing field considering PFs are relatively simple compared with
the numerous features that are offered through most DF systems. Thus, one of the primary quantitative aims of this paper was to fill this gap in the literature by examining what, if any, vocabulary gains are made when learners use VLSs together with PFs, in comparison to learners who use DFs. Qualitatively, the authors also wanted to systematically survey learners via a questionnaire to ascertain if the ubiquity, convenience and entertainment value of DFs are seen as advantages in the Japanese EFL context, as they were in other studies (e.g., Başoğlu & Akdemir, 2010). Once again, very limited research exists in the East Asian context, where populations have widespread Internet access and use of digital devices.

2. Literature review

2.1. Receptive vs. productive L2 vocabulary learning

Previous research has shown several differences between receptive vocabulary (RV), i.e., written or spoken words that a learner can understand (Burger & Chong, 2011) and productive vocabulary (PV), words learners can produce when they write or speak without external stimuli (Meara, 1999; Schmitt, 2000). The first is that a student’s RV is larger than her or his PV (Laufer, 1998; Fan, 2000; Webb, 2005, 2008). Additionally, studies have also been carried out to determine why such gaps exist. Laufer and Paribakht (1998) found EFL students, who had smaller RV - PV gaps compared to ESL students, benefited from learning differences, such as directly seeking out new words to use in authentic settings.

Another widely observed research outcome is that a learner’s RV improves faster than their PV. In a study over one year, Laufer (1998) found that a learner’s RV progressed very well, while little to no improvements were made in PV. Similar findings were made by Fan (2000), whose research indicated a slower rate of progress for PV. Again, Webb (2005) made similar observations; however, he also noted that learners with a larger RV were also more likely to know more PV. What was interesting with the Webb study is how the gains were achieved – via RV tasks (reading) or PV tasks (writing). He found that when equal time was spent on both tasks, RV proved more beneficial; however, when time was given according to the amount of time needed to complete a task, PV proved superior.

2.2. Digital flashcards vs. paper flashcards

Most comparison studies of DFs versus PFs have revealed that incorporating computer-assisted language learning (CALL) is more effective at enhancing L2 vocabulary learning. Başoğlu and Akdemir (2010) looked at the use of DFs on mobile phones versus PFs with a group of L2 English students at a Turkish university. While both groups were able to make significant gains, the DF group made greater significant improvements. In another study involving Turkish university students, Kiliçkaya and Krajka (2010) compared the usefulness of DFs via Wordchamp versus vocabulary notebooks and PFs. Not only did the DF group outperform the notebook and PF group on the post-test, but they also made greater significant gains on a delayed post-test. Azabaftari and Mozaheb (2012) also looked at the use of DFs with L2 English students at an Iranian university. The participants in the experimental group used a combination of DFs over mobile phones, short-message service (SMS) and the Internet to study the target vocabulary, while those in the control group used PFs. According to the post-test results, the use of mobile-learning and DFs had a greater positive effect on vocabulary learning than PFs. These studies demonstrate that DFs may help students better remember L2 vocabulary in the short-term, as well as support future recall to a greater degree than paper materials such as PFs and notebooks.

Despite the positive findings regarding DFs outlined above, not all comparison studies have resulted in superior gains by the DF group. Nikoopour and Kazemi’s (2014) study with university students in Iran had mixed results concerning the use of DFs to improve L2 vocabulary learning. Three learning methods were involved in the study: mobile phone flashcards, computer-based flashcards and PFs. While a significant difference was
not found between the paper and mobile phone flashcard groups, the computer-based group had significantly lower gains when compared to the PF group on a vocabulary post-test. The researchers posited that the ubiquity of mobile phones and PFs was the main reason for the discrepancy in vocabulary gains.

In regards to learner attitudes towards DFs to enhance vocabulary learning, there seems to be a preference for their use over PFs. The participants in Başoğlu and Akdemir’s (2010) research viewed DFs as the preferred method of studying vocabulary due to their efficacy, ubiquity, and entertainment value. Azabdaftari and Mozaheb (2012) found similar results in their research. They discovered three main benefits of DFs: ubiquity, convenience and vocabulary learning as entertainment. Similar to the aforementioned studies, Nikoopour and Kazemi (2014) also found that learners preferred DFs to PFs. On the other hand, there was no statistical difference in the learners’ opinions between the DF and PF group. Based on these findings, it seems that L2 students prefer the ubiquity and convenience of mobile learning with DFs to paper materials.

In short, not only do students seem to prefer DFs, but they have also been found to be more effective than traditional vocabulary learning methods. However, one limitation of the previously mentioned research is that the students using paper materials were not taught any VLSs to maximize the effectiveness of their learning. Given the positive effect that VLSs can have on L2 vocabulary acquisition (Kornell & Bjork, 2007; Mizumoto & Takeuchi, 2009), the lack of VLS instruction in those studies could have affected the results. Thus, the following research questions were investigated to fill this gap in the literature, as well as strengthen the limited empirical research of CALL-based flashcards in Japan:

- Were there any significant differences in receptive and productive L2 vocabulary improvements between a group of Japanese EFL learners which studied with PFs in conjunction with 3 VLSs, and another group which used DFs?
- How much time did each group spend studying the target vocabulary outside of class?
- What were the students’ opinions of each vocabulary learning method?

3. Methodology

3.1. Participants

The participants in this study were chosen via convenience sampling. They included a total of 52 1st-year EFL students at two private universities in Japan who were enrolled during the spring semester of 2016. The learners were divided into two equal groups: the PF group, which used PFs (n = 26), and the DF group, which used DFs (n = 26). The PF group was comprised of learners from only one of the universities, while the DF group was made up of students from both colleges. Students were placed in each group according to the availability of PCs in their respective classes. Vocabulary was a component of the grading criteria for all the classes involved; therefore, the use of PFs and DFs was an appropriate way to meet the specific needs of the students.

3.2. Target vocabulary

The New General Service List (NGSL) was designated as the target vocabulary for this study for several reasons. First, the NGSL provides learners with the most important high frequency words in English (Browne, 2013); thus, it can be considered as an essential component of L2 English learning. In addition, it offers a more modern and much larger corpus than its predecessor, the General Service List (GSL), which was developed more than 60 years ago by West (1953). Moreover, the NGSL affords learners more coverage with fewer words compared with the GSL. In other words, the NGSL provides students with a valuable resource to greatly expand their L2 vocabulary in an efficient manner without having to study multiple forms of a particular word (Browne, 2013).
3.3. Treatment

Aside from the differing study methods, both groups followed the same treatment procedure during the 12-week study. Each individual flashcard included the target vocabulary word and L1/L2 definitions. Students were given 15 minutes in class to study a new word list each week. The sole exception to this was the final week of the study, when the learners had the opportunity to review all the target vocabulary words. While studying outside of class was encouraged, it was not required nor was it tracked by the researchers.

The PF group was taught three VLSs to help enhance L2 vocabulary learning. Two of the VLSs, specifically oral rehearsal and association, were adopted from Mizumoto and Takeuchi’s (2009) research of VLSs with Japanese EFL learners. The researchers found that these two VLSs resulted in the greatest vocabulary improvements by the students, thus they were incorporated in the present study. The third VLS, dropping, was adopted from Kornell and Bjork’s (2007) research of flashcards. According to the researchers, dropping has the potential to promote memory recall, particularly if students do not have enough time to study. Because of this, dropping was seen as a useful strategy for the learners in the PF group. All three of the VLSs were taught and modeled prior to the start of the treatment. In addition, the VLSs were reviewed at the start of week two in order to reinforce the studying procedures.

The DF group studied via Quizlet and Cram, two popular online study tools. As of August of 2016, Cram had over 2.5 million members and more than 150 million user-created flashcards (Cram, 2016). Quizlet had even larger numbers – 40 million users per month with over 125 million user-created flashcard sets (Quizlet, 2016). Another important determining factor in choosing Quizlet and Cram is the fact that they are freely available for use on the web, in addition to offering free mobile apps through the iTunes’ App Store and the Google Play Store. As Bateson and Daniels (2012) note, the financial constraints of language instructors and students must be considered when incorporating CALL. Therefore, the results from this study have pedagogical implications to those who have limited financial resources and cannot afford paid or subscription-based vocabulary study services such as WordEngine and Anki.

Similar to the PF group, the students in the DF group received learner training to increase familiarity with the flashcard study tools before the start of the treatment. The researchers explained how to log in and use the specific features of each site. Tables 1 and 2 below show the features of the study tools in relation to RV and PV. Although the DF systems are similar, Quizlet offers slightly more PV tasks as well as additional corrective feedback based on the students’ responses. It is important to note that the participants were free to use Quizlet, Cram, or a combination of the two during the treatment. Furthermore, they were not encouraged to use one study tool over the other.

<table>
<thead>
<tr>
<th>Table 1. Features of Quizlet</th>
<th>RV activity</th>
<th>PV activity</th>
</tr>
</thead>
<tbody>
<tr>
<td>Word list</td>
<td>✔</td>
<td></td>
</tr>
<tr>
<td>Flashcards</td>
<td>✔</td>
<td></td>
</tr>
<tr>
<td>Test</td>
<td>✔</td>
<td>✔</td>
</tr>
<tr>
<td>Spell</td>
<td></td>
<td>✔</td>
</tr>
<tr>
<td>Learn</td>
<td></td>
<td>✔</td>
</tr>
</tbody>
</table>
Table 2. Features of Cram

<table>
<thead>
<tr>
<th>Feature</th>
<th>RV activity</th>
<th>PV activity</th>
</tr>
</thead>
<tbody>
<tr>
<td>Word list</td>
<td>✔</td>
<td></td>
</tr>
<tr>
<td>Flashcards</td>
<td></td>
<td>✔</td>
</tr>
<tr>
<td>Test</td>
<td>✔</td>
<td>✔</td>
</tr>
<tr>
<td>Memorize</td>
<td></td>
<td>✔</td>
</tr>
<tr>
<td>Jewels of Wisdom game</td>
<td>✔</td>
<td></td>
</tr>
<tr>
<td>Stellar Speller game</td>
<td></td>
<td>✔</td>
</tr>
</tbody>
</table>

Not only do Quizlet and Cram offer both RV and PV learning activities, but these features also involve specific types of VLSs to help promote vocabulary recall. In a study of L2 VLSs, Lawson and Hogben (1996) categorized several types of strategies that learners used, two of which are relevant to the use of DFs in this study: repetition and word feature analysis. Tables 3 and 4 below detail the different forms of each VLS (p. 115). It is important to note that all of the strategies listed in the tables were incorporated in at least one or more of the features that Quizlet and Cram offered.

Table 3. Repetition strategies

<table>
<thead>
<tr>
<th>Forms of repetition</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Reading of Related Words</strong></td>
</tr>
<tr>
<td>The student reads words that are related to the meaning of the word, e.g., L1/L2 definitions of the target word.</td>
</tr>
<tr>
<td><strong>Simple Word Rehearsal</strong></td>
</tr>
<tr>
<td>The student repeats the word and/or its meaning. In the case of DFs, this can be done by the online tools themselves as a listening activity.</td>
</tr>
<tr>
<td><strong>Writing Word and Meaning</strong></td>
</tr>
<tr>
<td>The student writes out the target word and/or its definition.</td>
</tr>
<tr>
<td><strong>Cumulative Rehearsal</strong></td>
</tr>
<tr>
<td>The student reviews words that were previously studied.</td>
</tr>
<tr>
<td><strong>Testing</strong></td>
</tr>
<tr>
<td>The student tests himself/herself by generating either the target word or its definition.</td>
</tr>
</tbody>
</table>

Table 4. Word feature analysis strategies

<table>
<thead>
<tr>
<th>Forms of word feature analysis</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Spelling</strong></td>
</tr>
<tr>
<td>The student spells out the word.</td>
</tr>
</tbody>
</table>
3.4. Data collection and analysis

Quantitative data was obtained via vocabulary assessments administered at the outset and completion of the treatment (Appendices 1 and 2). The tests were administered to determine which study method – PFs or DFs – was more effective at enhancing the two dependent variables: 1) RV knowledge of the NGSL, and 2) PV knowledge of the NGSL. The RV assessment, which was modeled after the Vocabulary Size Test (Nation & Beglar, 2007), was developed by Stoeckel and Bennett (2015). Because there was no PV test of the NGSL, the researchers created a PV assessment which adapted the same format as the productive vocabulary levels test by Laufer and Nation (1999). For each target word, a meaningful context sentence was provided as well as the first few letters of the term in order to eliminate any other possible answers. The test items were chosen randomly from levels 1 and 2 of the NGSL. Correct answers on the RV test were excluded from the PV assessment to ensure that there were no duplicate responses. Initially, the participants took tests based on the first two word levels of the NGSL. However, both the PF and DF group did not show mastery, i.e., a score of at least 80–85%, at either level of the RV test (Stoeckel & Bennett, 2015). Therefore, the groups studied level 1 of the NGSL during the 12-week treatment.

Qualitative data was collected through L1 surveys which were administered after the completion of the post-test. The first question on each questionnaire asked the students to report their estimated vocabulary study time outside of class. Items two through nine were comprised of Likert-type questions asking the students to rate their views towards the study methods according to a scale ranging from strongly disagree (1) to strongly agree (5). The final question related to the study preferences of each method, namely, which VLS the PF group found most useful and the preferred DF study tool for the experimental group.

4. Results

An independent t-test was used to compare the pre- and post-test results between the two groups. However, it is important to note that the groups were found to be unequal, i.e., there was a significant difference in RV between the PF group (M = 12.65, SD = 4.1) and the DF group (M = 15.38, SD = 3.13) at the beginning of the treatment, t(50) = 2.7, p = .009. Similarly, differences were found in PV with the PF group (M = 6.96, SD = 2.31) producing significantly lower scores than the DF group (M = 8.77, SD = 2.75) t(50) = 2.5, p = .013. A paired t-test was used to analyze gains made within each group. Descriptive statistics are also provided to show the vocabulary improvements from the pre- to post-test as well as the results from the post-treatment survey.

<table>
<thead>
<tr>
<th>Table 5. Pre-test and post-test results</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
</tr>
<tr>
<td><strong>Pre-test</strong></td>
</tr>
<tr>
<td>Receptive</td>
</tr>
<tr>
<td>12.65</td>
</tr>
<tr>
<td>Productive</td>
</tr>
<tr>
<td>6.96</td>
</tr>
<tr>
<td><strong>Post-test</strong></td>
</tr>
<tr>
<td>Receptive</td>
</tr>
<tr>
<td>15.19</td>
</tr>
<tr>
<td>Productive</td>
</tr>
<tr>
<td>10.08</td>
</tr>
</tbody>
</table>

Table 5 above shows the results of the pre- and post-tests of each group. A paired t-test indicated that the PF group was able to make a significant improvement in RV from the pre-test (M = 12.65, SD = 4.10) to the post-test (M = 15.19, SD = 2.74), t(25) = 2.81, p = .009. The DF group also had a significant increase in RV from pre-test (M = 15.38, SD = 3.13) to the post-test (M = 17.85, SD = 2.36), t(25) = 4.15, p = .0003. While both groups were able to make significant gains, an independent t-test revealed the gains were not significantly different between the PF group (M = 2.54, 4.61) and the
DF group (M = 2.46, SD. 3.02), t(50) = .071, p = .94. Similar results were also found in relation to PV improvements. A paired t-test showed that the PF group made significant gains in PV from the pre-test (M = 6.96, SD = 2.31) to the post-test (M = 10.08, SD = 2.23), t(25) = 4.86, p = .0001. The DF group made significant improvements as well from the pre-test (M = 8.77, SD = 2.75) to the post-test (M = 12.08, SD = 2.17), t(25) = 7.4, p = .0001. However, there was not a significant difference between the PV increase between the PF group (M = 3.12, SD = 3.27) and the DF group (M= 3.31, SD = 2.28), t(50), = .24, p = .81.

![Figure 1. Amount of time the PF group studied target vocabulary outside of class.](image1)

![Figure 2. Amount of time the DF group studied target vocabulary outside of class.](image2)

Figures 1 and 2 above show a breakdown of the amount of time groups spent studying the target vocabulary outside of class. A much larger percentage of the students in the PF group (62%) took advantage of the opportunity to study the words outside of class compared with the DF group (31%). In fact, more than two-thirds of those studying with DFs did not study the vocabulary at all, which is nearly two times more than those who chose not to study with PFs.
Table 6. Percentage of agreement towards survey statements

<table>
<thead>
<tr>
<th>#</th>
<th>Statement</th>
<th>PF group</th>
<th>DF group</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>I was able to learn English vocabulary more quickly with paper/digital flashcards.</td>
<td>50.0%</td>
<td>53.8%</td>
</tr>
<tr>
<td>2</td>
<td>Using paper/digital flashcards improved my English vocabulary.</td>
<td>65.3%</td>
<td>57.9%</td>
</tr>
<tr>
<td>3</td>
<td>Using paper/digital flashcards made it easier to learn English vocabulary.</td>
<td>61.5%</td>
<td>65.3%</td>
</tr>
<tr>
<td>4</td>
<td>I think paper/digital flashcards were useful in my class.</td>
<td>53.8%</td>
<td>69.2%</td>
</tr>
<tr>
<td>5</td>
<td>It was easy for me to study English vocabulary with paper/digital flashcards.</td>
<td>50.0%</td>
<td>69.2%</td>
</tr>
<tr>
<td>6</td>
<td>It was easy for me to become skillful at studying English vocabulary with paper/digital flashcards.</td>
<td>61.5%</td>
<td>69.2%</td>
</tr>
<tr>
<td>7</td>
<td>Learning how to study English vocabulary with paper/digital flashcards was easy for me.</td>
<td>42.3%</td>
<td>73.0%</td>
</tr>
<tr>
<td>8</td>
<td>I prefer studying English vocabulary with paper/digital flashcards to digital/paper flashcards.</td>
<td>53.8%</td>
<td>69.2%</td>
</tr>
</tbody>
</table>

Table 6 shows the percentage of agreement towards the statements on the questionnaire. Besides statement two, “Using paper/digital flashcards improved my English vocabulary,” there were higher levels of agreement in the DF group. In particular, items four, five, seven, and eight illustrated more favorable views towards DFs, with these statements receiving at least 15 percent or more agreement in the DF group than the PF group. Overall however, opinions of both PFs and DFs were generally positive, with only one statement receiving lower than 50% agreement (PF group, item seven).

![Figure 3. Perceived usefulness of the vocabulary learning strategies.](image-url)
Figures 3 and 4 show the study preferences of each group, i.e., the VLS which had the highest percentage of perceived usefulness as well as the preferred DF system. Over half of the participants in the DF group stated that they found oral rehearsal to be most useful (58%). This was followed by association (31%) and dropping (11%). In terms of DFs, there was a slight preference for Quizlet over Cram among the students in that group, with more than half of them selecting the former as their preferred study tool.

5. Discussion

While both the PF and DF group were able to make improvements in receptive and productive vocabulary, the gains made within each group did not significantly differ. This is in contrast to previous research (Azabdaftari & Mozaheb, 2012; Başoğlu & Akdemir, 2010; Kılıçkaya & Krajka, 2010) concluding that traditional forms of vocabulary learning, including PFs, were not as effective as DFs. However, as aforementioned, the students using PFs in these studies were not taught any VLSs. While DF systems often provide learners with a variety of ways to study, PFs are a much more basic form of vocabulary learning. Therefore, students ought to be taught VLSs in order to maximize the effectiveness of PFs to retain new vocabulary. Another interesting finding based on these results is that learners’ PV can improve at the same pace as their RV. This differs from previous research which showed that L2 students made little to no progress in PV (Fan, 2000; Laufer, 1998; Webb, 2008).

While the ubiquity and convenience of DFs seem to make them an appealing vocabulary learning method for students (Azabdaftari & Mozaheb, 2012; Başoğlu & Akdemir, 2010), students in the DF group reported much lower levels of vocabulary study time. These results indicate that these variables may not be as influential in a student’s decision to study a L2 outside of class. Factors such as learner motivation, attitudes towards the target language, as well as other external variables may play a greater role in the amount of time a learner chooses to study.

The participants’ responses to the survey signify that the DF group preferred their method of studying to a greater degree than the PF group. Notably, ease of use seems to be a distinct advantage of DFs over PFs, as shown by the high levels of agreement towards statements five through seven. The ability to study DFs anytime and anywhere via smartphone may have contributed to these results, whereas PFs are much more inconvenient to use on the go, thereby decreasing their value. These findings indicate that the DF group was more satisfied with their method of vocabulary study and as a result, would be less likely to switch to PFs if given the opportunity.
Out of the three VLSs taught to the participants in the PF group, oral rehearsal was found to be the most useful strategy in the eyes of the learners. Association was also perceived to be useful by a significant percentage of the students (31%). These results coincide with the findings of Mizumoto and Takeuchi (2009) that showed Japanese L2 English learners had favorable views towards the perceived usefulness of oral rehearsal and association to improve vocabulary learning, as well as the research of Kornell and Bjork (2007), which showed that it may be difficult to effectively utilize the strategy of dropping in vocabulary learning. Although DF systems have built-in features that promote the use of VLSs, PFs are much more simplistic; therefore, teachers ought to encourage the use of VLSs in order for L2 students to maximize their vocabulary development, especially with paper materials.

In terms of DFs, there was a slight preference in favor of the use of Quizlet. While Cram has similar features, the fact that Quizlet offers more PV activities as well as additional corrective feedback may have led to greater interest in the program. Another possible explanation for the preference may be due to the games, as they have been positively received in a previous study which incorporated Quizlet (Jackson III, 2015). Despite these findings, more research needs to be done in order to compare not only the views L2 learners have towards different DF systems, but also the potential vocabulary improvements that can be made.

6. Conclusion

One of the primary goals of this study was to compare the effectiveness of DFs and PFs to improve RV and PV knowledge in a L2. In this regard, both methods have been found to be equally as effective, which goes against most previous research on the topic of L2 vocabulary learning and DFs (Azabdaftari & Mozaheb, 2012; Başoğlu & Akdemir, 2010; Kılıçkaya & Krajka, 2010). Unlike those studies however, the present study incorporated VLSs in the PF group to help promote vocabulary learning due to the simplistic nature of the study tool when compared with more sophisticated DF systems. These findings highlight the importance of VLSs as a way to enhance vocabulary learning with PFs when DFs are not a viable option. In terms of learner opinions, the participants in this study preferred DFs over PFs, with ease of use being one of the key factors. Nevertheless, the PF group spent more time studying the target vocabulary outside of class, indicating that the advantages inherent to DFs may not be enough to motivate students to study in their own time. Thus, language teachers must stress the benefits of vocabulary learning and encourage students to take full advantage of any opportunities to study the target language outside of class, regardless of whether or not it’s a CALL-based activity.

One of the limitations of this study is the fact that the groups were not equivalent as it pertains to the dependent variables. This was unavoidable due to the incorporation of convenience sampling. Future studies should examine these variables with homogenous groups of students who are chosen via random sampling. Additionally, it is not known if external variables affected the results of the study. Therefore, future research ought to administer a pre-treatment survey to take into account other factors such as smartphone ownership or Internet access. Furthermore, it is unclear how much of an impact the VLSs had in supporting L2 vocabulary development in the PF group. Thus, it may be worthwhile to employ a future study with a PF group, a PF & VLS group, and a DF group in order to understand how much of a role VLSs played in enhancing vocabulary. Lastly, although Quizlet and Cram are comparable in terms of their features, it would be interesting to compare the efficacy of the two DF systems to improve L2 vocabulary among L2 students.

References


Appendix 1
Level 1 of the Test of Written Receptive Knowledge of the New General Service List. (Stoeckel and Bennett, 2015).

1 charge: They are the charges.
   a. important things to think about
   b. prices for a service
   c. good things
   d. reasons

6 policy: That is a good policy.
   a. kind of school
   b. story
   c. place to visit
   d. way to act

2 case: This is a good case.
   a. place to study
   b. way something works
   c. example of something
   d. plan for the future

7 rise: They will rise next week.
   a. become higher
   b. change
   c. become better
   d. finish

3 different: They are different.
   a. easy to see
   b. large
   c. not easy
   d. not the same

8 sure: I am sure.
   a. young
   b. early
   c. certain
   d. new

4 room: Where is the room?
   a. thing we read
   b. thing to drive
   c. place to buy things
   d. space in a building

9 health: Health is important.
   a. learning in a school or college
   b. having no problems with your body
   c. learning by doing something a lot
   d. having help from other people

5 lead: I will lead you.
   a. take you to a place
   b. meet you
   c. let you
   d. give something to you

10 expect: I expected this.
   a. thought this would happen
   b. said this idea
   c. put this into something
   d. took this to a place
11 include: We are including it.
a. paying
b. changing
c. adding
d. reading

16 return: Please return it.
a. talk about it
b. sell it
c. show it
d. take it back

12 building: Where is the building?
a. group of people working together
b. road
c. middle part
d. place to live or work

17 result: We had the same results.
a. questions
b. thoughts
c. rules for doing something
d. things that happened at the end

13 true: that is true.
a. correct
b. different
c. interesting
d. natural

18 among: He was among them.
a. after
b. behind
c. together with
d. not far from

14 teacher: They are teachers.
a. people with children
b. workers in schools
c. leaders in a company
d. young people

19 consider: She considered it.
a. could not find
b. needed
c. thought about
d. said

15 well: You did that well.
a. fast
b. in a good way
c. by yourself
d. often

20 approach: We like your approach.
a. way of doing something
b. part of a book
c. house and land
d. facts and information

Appendix 2
Level 1 of the Test of Written Productive Knowledge of the New General Service List.

Complete the underlined words. The example has been done for you.

He was riding a bicycle.

1. Bo_______ my mother and father are teachers.
2. Please sh_______ her how to use the computer.
3. The children sat in the ce_______ of the room.
4. I like to take pic_______ of my family and friends.
5. You need to p_______ for these movie tickets.
6. I left my phone at home. I think i_______ is on the kitchen table.
7. He is a baker. He ma_______ bread.
8. She works at an org_______ which helps poor children.
9. Students must obey many ru_______ at school.
10. He would like to tr_______ around the world.
11. I don't und_______ what he is saying.
12. I'm looking for my sunglasses. Did you see th_______?
13. The weather is very b_______ today. It's raining hard.
14. She fol_______ him into the house.
15. John and Alice are a nice cou_______.
Research and development

A freely-available authoring system for browser-based CALL with speech recognition

Myles O’Brien
Mie Prefectural College of Nursing, Japan

myles.obrien@mcn.ac.jp

Abstract
A system for authoring browser-based CALL material incorporating Google speech recognition has been developed and made freely available for download. The system provides a teacher with a simple way to set up CALL material, including an optional image, sound or video, which will elicit spoken (and/or typed) answers from the user and check them against a list of specified permitted answers, giving feedback with hints when necessary. The teacher needs no HTML or Javascript expertise, just the facilities and ability to edit text files and upload to the internet. The structure and functioning of the system are explained in detail, and some suggestions are given for practical use. Finally, some of its limitations are described.

Keywords: Automatic speech recognition, CALL authoring tool, computer-assisted language learning, Google speech API.

1. Introduction
The quality of automatic speech recognition (ASR) has been improving steadily with technological advances. The history of ASR-based CALL dates back to the end of the last century (Aist, 1999; Bernstein, Najmi, & Esani, 1999; Strick, 2012, p. 10), when there were experiments with prototype systems, notably FLUENCY (Eskenazi, 1999), a computer-assisted pronunciation training (CAPT) system which used carefully constructed output, free of explicit prompting, to elicit an oral response which was confined to a very narrow range of possibilities. But progress in this area has been rapid, and there are now several commercial systems (Witt, 2012, p. 5) which can evaluate speech with a good correlation to the judgement of human assessors (Bernstein, Van Moere, & Cheng, 2010; Zechner et al., 2014). These, and other systems under development (Penning de Vries, Cucchiarini, Bodnar, Strik, & van Hout, 2015; van Doremalen, Boves, Colpaert, Cucchiarini, & Strik, 2016) attempt to evaluate grammar and content in addition to pronunciation.

The uses of ASR are not confined to CALL, of course. One company which has been developing ASR for its own various purposes, and is among the leaders in the field, is Google. The ability to perform a Google search by voice alone became available on iPhone in 2008. It has since been greatly extended and improved, now being available in over 80 languages. Google fully opened its Speech API, which gives direct access to its speech recognition capabilities, to developers for use in their applications on a commercial basis in 2016 (https://cloud.google.com/speech). However, a free tier of indirect access to basic ASR functions using JavaScript calls from the Google Chrome browser has been available since the release of version 25 in 2013. This enables speech-to-text conversion within the browser, and so offers the interesting possibility of browser-based CALL incorporating ASR. In order to implement this in as useful a way as
possible, it was decided by the author to attempt the development of a very flexible system which could be used with ease by anybody, without JavaScript knowledge, to make their own ASR-based CALL material for internet deployment. The system has been successfully implemented and made freely available, and is described in detail in this paper.

2. Outline of the system

The system is tentatively named QApeak. At its most basic, it consists of one html file (and a folder containing the image files it uses) plus a plain text file containing a list of questions with the acceptable answer(s) for each. The person making the CALL material (the teacher) needs to edit only the text file. The person using the material for study (the user) can elect to listen to and/or read the question, and answer by speaking or typing. The teacher has the option of adding a media file (image, sound, or video) and/or text of any length to each question. Another option is to add a sound file of the question. If this is not included, the question text will be read by the device’s text-to-speech function. Figure 1 shows an example of the interface for a question which includes both image and text options.

![Figure 1. An example of the interface for a question.](image)

The user hears "What's the dog doing?" and is expected to provide an appropriate response, like "It's catching the ball." The user interacts through the icons, which are in the black strip, and the answer box, which is just below it. Figure 2 shows an annotated version of the same interface.
Leftmost in the black strip is the progress indicator, which shows how far the user has advanced through the set of questions. Next is the ear icon, which allows the user to hear the question through text-to-speech, or a sound file, if available. Then the microphone icon activates Google speech recognition for the user to speak the answer. The ASR system’s interpretation of the speech appears in the pink answer box, and the answer is checked against the list of acceptable alternatives the teacher has set up. If the answer is correct, a congratulatory image and a green arrow icon, to move on to the next question, appear. Also, the other remaining permitted answers are displayed. If the answer is incorrect, corrective feedback (described in detail later) is displayed, and the user can try speaking the full answer again, or edit the current text in the answer box, hitting the “Enter” key on the device to have the answer checked. This process can be repeated until a correct answer is obtained, or the user resorts to the “Give Up” button, which shows the full list of permitted answers and allows progress to the next question. One more source of help is available along the way: hitting the eye icon will display the text of the question at any stage, which may assist a user who is having trouble understanding the audio of the question. The user is also permitted to type their answer directly at any stage, even from the beginning without speaking at all.

3. Anatomy of the controlling text file

To set up the questions, the teacher needs to make a plain text file, with a very simple format, specifying the questions, permitted answers, and additional media files. Each question-answer set extends over 3 or more lines. The first line must begin with a question mark. This specifies the beginning of a new question. If a media file is to be included, its name is entered after the question mark, and optional text may be typed on the same line. The next line should contain the text of the question, and the sound file name, if one is supplied, to be used instead of text-to-speech. The third and subsequent lines contain the permitted answers. So, in general, the format of a question is like this (where square brackets signify an optional item):

? [media file name] [text]
Question text [mp3 file name]
Answer 1
Answer 2
Answer 3
etc.
The example shown in Figure 1 corresponds to the following lines:

? dog.jpg Note: this is a female dog
What's the dog doing?
She's catching a ball.
It's catching a ball.

Lines where the first character is not “?” or alphabetic (including blank lines) are ignored, so that dividers or comment lines can be added anywhere. Any number of questions may be included in one text file.

These are the first 3 questions from the text file for the online example (http://www.mcn-moodle.org/asr), which shows that the format is very simple, yet flexible:

? What browser does this have to be?
Chrome
It has to be Chrome.

? Jack can't stand carrots.
Does Jack like carrots?
No, he doesn't.
------XXXXXXXXXX-------

? This is my sister's cat. His name is Nando. nando6.mp4
What is he swinging?
His tail.
He's swinging his tail.
He is swinging his tail.

The first is a minimal example with no media files at all and just 2 permitted answers. The second adds a little optional explanation. It accepts only one answer. “------XXXXXXXXXX-------” will be ignored. The third has optional text and a video file.

While direct editing of the text file affords the advantages of maximum speed, simplicity, and flexibility, the procedure may not appear very user-friendly to many less computer-oriented teachers, and could well discourage some from using the system at all. Therefore, if the initial version of the system attracts attention and proves successful, a high-priority addition to the next version should be a form-like application to enable structured, guided input of the text items and file names, and automatic generation of the corresponding controlling text file.

4. Uploading to the web

The html file and “images” folder which are supplied as the core of the system, the controlling text file, and all specified media files should be uploaded to the same directory. The html file may be freely renamed, but the controlling text file must have the same name, with a “.txt” extension in place of “.html”. Figure 3 shows a schematic example:
Using this system, deploying ASR-based CALL material to the web requires minimal technical expertise from the teacher. The required files just need to be dropped into a folder which is already online, or which will then be uploaded. How well the material which can be made with the system meets the requirements of each teacher is, of course, a separate question. There is sufficient flexibility in the system to allow it to be used in quite a variety of different ways, some of which will be outlined in a later section. Before that, the description of the functioning of the system will be completed, with an account of the feedback it produces.

5. Feedback

When the user has finished speaking or hits the “enter” key to have an answer checked, the system first checks if the user’s answer matches any of the teacher’s listed permissible answers exactly. If it does, a message appears, stating how many attempts were needed, and whether the user looked at the question text or did some typing. Also, the full list of permitted answers for that question is displayed, and the arrow icon to allow progression to the next question appears. Finally, one of four images appears, depending on how smoothly the user has arrived at the answer. The user incurs a penalty for: (1) looking at the question text, (2) typing an answer completely or in part, or (3) requiring more than one attempt to answer correctly. The 4 images supplied with the system are shown in Figure 4. These are kept in the “images” folder, and the teacher may replace any or all of them with different .jpg images of the same filename. The image score3.jpg appears when the user incurs no penalties, i.e., gives correct answer first time by listening and speaking only. The other images, score2.jpg, score1.jpg and score0.jpg appear when 1, 2, or 3 penalties, respectively, are incurred. Note that extra penalties are not incurred for repeated “offences” of the same type, so the maximum is 3.
Figure 4. The supplied feedback images.

When the user’s answer does not match exactly any of the designated correct answers, corrective feedback appears. If the length of the user’s answer is too far from the length of any of the correct answers (word count less than 0.6 times or greater than 1.5 times the word count of the correct answer), a message appears saying this, and the user must try again. Otherwise, feedback is based on the correct answer which most closely matches the user’s attempt, and shows the words which correspond exactly in both, in their correct positions. Incorrect words are substituted by a number indicating the number of letters in the correct word. To allow for a sequence of correct words in the user’s answer which is slightly offset as regards position in the sentence, it is compared to the model answers in five alignments, with offsets ranging from 2 words to the left to 2 to the right, to find the longest matching sequence of words (best match). Figure 5 illustrates the process.

Figure 5. An example of checking the user’s answer in the 5 alignments.
A couple of examples are now given, for further explanation:

(1) Correct answer: They’ve gone to the movies.
User’s answer: They have gone to the movies.
Feedback: 7 gone to the movies.

Note that the user’s answer is correct as regards grammar and meaning, but if it is not listed among the teacher’s permitted answers, it is regarded as wrong. In this case, the teacher has decided to enforce the use of a contracted form. “gone to the movies” is correct, so it is displayed in full, and “7” signifies “They’ve”, which has seven characters including the apostrophe.

(2) Correct answer: He went there yesterday.
User’s answer: He went to there yesterday.
Feedback: He went 5 9.

Note that there are 2 correct word sequences of 2 words each in the user’s answer, but only one is shown. Showing both would give away the complete correct answer in this case. Of course, if the student’s answer had been “She went to there yesterday”, the feedback would have been “2 4 there yesterday”.

This feedback is intended to provide the user with a solid basis to improve the next attempt, without making the correct answer too obvious. In fact, disclosure of the number of letters in each missing word will narrow down the possibilities considerably and, in many cases, allow the user to answer next time with a high degree of confidence. It might be argued that this makes the challenge too easy, and displaying something like “BLANK” for each word instead of the number of letters would be preferable. This would correspond to the “indirect CF” of Ellis (2009), specifically his type 2a, “indicating + locating the error” (p. 98), whereas the method used could be said to fall between this and his “direct CF”, where the correct answer is provided. However, if the user does find that the feedback makes the correct answer very easy to determine, this probably means that they already have quite a good idea what it should be, from a narrow range of possibilities. Also, an additional factor which should be considered with this system is that, where the user is trying to answer by speaking, even if the correct answer is known, an additional challenge is to pronounce it aloud so that Google ASR will recognize it as intended. This factor swung the decision in favor of using what we might call “semi-direct CF” instead of indirect CF in this original version of this system. A variant version would, of course, be very simple to produce. As the system allows “hybrid” input by speaking and/or writing, and also can be used in a variety of different ways, as will be discussed in the next section, it is difficult to specify an optimum one-size-fits-all type of feedback. It is all the more difficult since, even in a given situation, such as pronunciation (Golonka, Bowles, Frank, Richardson, & Freynik, 2014, p.81-82; Levy & Stockwell, 2006, p189-190) or composition (Ellis, 2009; Guenette, 2007) training, the question of what kind of CF (including none) is most effective remains contentious.

6. Usage possibilities

Basically, the system provides the teacher with a simple way to set up browser-based CALL material which will elicit spoken or written answers from the user and check them against a list of specified permitted answers, giving feedback with hints when necessary.

Typically, the response will be elicited through a direct question, which the user can listen to and/or read, and the user may need to refer to the image/sound/video and/or extra text, if provided. In the simplest case (e.g., “What’s the day after Tuesday?” to elicit “It’s Wednesday”), no extra information is required. In the example shown in Figure 1, the user will need to refer to the image to answer the question “What’s the dog doing?” Straightforward reading or listening comprehension questions may be
A possibility that doesn’t use direct questions to elicit the response is “listen and repeat”, which combines listening and pronunciation practice. The text could also be shown, to remove the listening comprehension element and concentrate on pronunciation, though the sound file would still serve as a pronunciation model. Conversely, the user could be required to type the answer to put the focus on listening, making a classic dictation exercise. Of course this system could not provide very rigorous CAPT, since the feedback is entirely based on the results of Google’s ASR, which uses AI techniques to make sense of the input by combining results from its Acoustic, Pronunciation, and Language models (https://www.google.com/about/careers/stories/how-one-team-turned-the-dream-of-speech-recognition-into-a-reality), so that intelligible, rather than native-like, pronunciation (Munro, 2011; Witt, 2012) is sufficient to produce a correct result.

The system allows a teacher to harness the power of Google’s ASR in a very flexible and simple way, though it has several limitations.

### 7. Limitations

The system works well on Windows, Macintosh, and Linux desktop or laptop computers with a fast internet connection, but the Google Chrome browser must be used. ASR does not work with other browsers, though all the other features do. Although it is designed to adapt to all screen sizes, the system’s ASR will not work at all on iOS devices, even if Chrome is used. It does work on Android, but the ASR results can be much inferior to those obtained on desktop or laptop machines. Even under the best conditions, Google ASR is, of course, not perfect. It may be very difficult for a speaker to get it to successfully interpret a short, single word like “two”, as there are so many variations in how individual speakers may pronounce it, and there is no context supplied to aid interpretation. However, a phrase like “one, two, three” will be recognized far more easily. Similarly, unless it is pronounced in a very specific way, the single word “lung” will be taken as the much more common word “long”, but “heart, lung, kidney” causes no such problem. Of course, much the same would apply to a human listener. The teacher should bear this in mind when designing material.

For maximum compatibility across platforms, image files are limited to the three types, jpg, png, and gif; sound files must be mp3, and video files mp4. When suitable media files are available, they are reasonably easy to deploy by just including their filenames at the appropriate points in the setup file and uploading the files themselves with the other system files. However, the teacher must be careful to avoid copyright breaches, and making customized media files is time-consuming and requires some technical expertise. Each media file must be separately included. There is no way of, for instance, using different sections of a large sound or video file.

The teacher needs to have access to a website to which they can freely upload files, and the minimal expertise to do so. If HTTPS protocol is not used, each time the user hits the microphone icon to input speech, they may get an additional screen message requiring confirmation of permission to use the microphone, and it will not be activated until they hit the Allow button. Recent versions of the Chrome browser seem to have become much less intrusive in this regard, requiring confirmation only once at the start of a session. Also, HTTPS support is becoming more common as a free option on hosting services. If it is not available by default, an SSL certificate to enable it can be obtained from a Certificate Authority such as Let’s Encrypt, https://letsencrypt.org, which is a non-commercial organization offering a free, automated service.

The teacher’s model answers may contain punctuation and capitalization, but because of the difficulty of including these through ASR, they will be ignored in checking the user’s answer.
8. Pilot study

A small pilot study to gauge user reaction was carried out individually with six Japanese nursing students, of mixed English ability, whose course includes several compulsory English subjects. The functioning of the system was explained with a few examples, and then they were asked to try it out, using a set of 25 questions, a few very simple (Are you a student? → “Yes, I am”) and others a little more challenging (being shown a picture of a nurse with a stethoscope around her neck, and asked What does she have around her neck? → “She has a stethoscope”), but none were very difficult, in order to avoid the difficulty of the material itself distracting from the central consideration of the principles of the system’s functioning and its interface. When they had finished all the questions, they were asked to complete a short web questionnaire in Japanese consisting of three 5-point scale items, (in English translation) “Is likely to help your study of English?”, “Is easy to use?”, “Is enjoyable?”, and three free input items, “Good points”, “Bad points”, “Any other comments”. The researcher observed their use of the system, providing assistance if needed, but the questionnaires were completed in private, to avoid pressure on the students.

The responses for the 5-point scale items were very positive, an average of 4.5 for “Is easy to use” and 4.8 for the other two items. In the free input section, positive points mentioned were, “helpful for pronunciation training” (most frequent), ”enjoyable to use”, “alternative answers and hints shown”, and “can enjoy studying even if weak at English”. Negative points or suggestions for improvement included desires for an enhanced hints option showing what words should be used (not just how many letters in each) and allowing their pronunciation to be heard, better microphone sensitivity (low voices were not picked up well), and a greater range of acceptable answers (though this last point is not inherent to the system, but at the teacher’s discretion).

Observation of the students’ trial suggested that the system has the potential to be a useful tool for language study. Some common shortcomings of Japanese speakers’ pronunciation of English were apparent in particular items, for instance words with the “l” sound, like “yellow” and “cold”, or the word “would”, which the weaker students tend to pronounce as “ud”. The Google ASR interpretation was often absurdly far from the intended utterance (for example, an intended “It’s yellow” was interpreted as “8 year old” for several of the students), but they were very pleased if they could eventually get their intended words across after several attempts. In the worst cases they resorted to typing, but nobody in the trial went as far as using the “Give up” button.

The overall reaction was quite positive, indicating that the system may have a lot of potential as a tool for the language teacher. It is hoped that it will be applied and prove beneficial in many different teaching environments, and that improved or variant versions will make it all the more useful and adaptable.

9. Download

The system can be downloaded from http://www.mcn-moodle.org/asr. The zip file includes the HTML file and images folder, which contains all the image files used in the user interface. These may left unchanged, though teachers are free to make their own customizations by editing the HTML file or replacing any image files with their own. The controlling text file used in the online example is also included. The teacher will need to edit this, or make a new one, to set up new material. The author also gives permission for teachers to make modified versions of the system for educational use by editing the HTML or JavaScript, provided they do not claim the original or modified system as their own work.
References


Reflective practice

Using Facebook to improve L2 German students’ socio-pragmatic skills

Axel Harting
Institute for Foreign Language Research and Education, Hiroshima University, Japan

Abstract
This study explores ways of using Facebook as a tool to improve the pragmatic competence of students of German as a foreign language in Japan. Nine students of a textbook-based German class (CFER level A2) voluntarily participated in this blended learning approach, in which they were assigned weekly online tasks aimed at eliciting speech acts commonly used in online interaction. The tasks required the students write posts concerning their daily routines onto the timeline of a dedicated Facebook Group page and to comment on each other’s posts. In order to find out what difficulties learners face when producing certain speech acts, the students’ posts and comments were analysed qualitatively and quantitatively by determining the frequency, accuracy, and appropriateness of speech acts performed per task. The results suggest that the tasks developed seemed to be appropriate for eliciting a large number of speech acts, while types, frequency, and appropriateness of the speech acts produced varied significantly. As was revealed, difficulties in task performance did not only stem from a lack of L2 (socio-pragmatic) knowledge, but also from inexperience in the use of the network itself. For task performance, students strongly relied on the expressions provided in pre-task activities. As the results of a post-course students’ survey indicate, most students agreed that using Social Networking Sites (SNS) is an appropriate tool for language learning, and that they were able to increase their knowledge and use of German speech acts significantly during this course. However, only those students who frequently used SNS in their everyday lives seemed to be able to reap the full benefits of the project.

Keywords: Facebook, Social Networking Sites, pragmatic competence, speech acts, blended learning, collaborative learning.

1. Introduction
This paper is based on the commonly held assumption that Japanese L2 learners lack pragmatic competence in German as a foreign (GFL) language as well as opportunities for authentic L2 interaction (Harting, 2006 and 2008). In order to enable students to tap into online and offline German resources through SNS (Social Networking Sites) and at the same time to provide opportunities to use the L2 outside the classroom, the author introduced a Facebook-project in a third-year textbook-based GFL class, which addresses the following research questions:

1. How do GFL students in Japan respond to using Facebook for learning German?
2. How should Facebook tasks be designed to facilitate the acquisition of speech acts?
3. Is Facebook an appropriate tool to improve GFL students’ socio-pragmatic competence?

The analysis is based on the textual communication of nine Japanese university learners of German, adjudged by their German teacher to be at the A2 CEFR level. In order to analyse authentic communication, the study employs Facebook posts and comments. Weekly tasks related to the learners’ daily routines were provided and students were asked to make their own posts and respond to their classmates’ posts. These subsequent responses offered opportunities for the learners to engage in a variety of speech acts (e.g. greetings, thanks, apologies, requests, as well as expressions of (dis-)like, (dis-)agreement, surprise, and emotion). The posts and comments on the Group’s timeline were analysed quantitatively and qualitatively according to the frequency, accuracy, and appropriateness of speech acts produced per task. Class discussions ensued, and evolved organically to consider the relative pragmatic appropriateness. Such discussion appears to raise both pragmatic competence as well as awareness. Results and implications of this explorative case study are discussed in this paper, along with potential suggestions for future practice and research.

2. Theoretical Background

Due to the growing popularity of SNS such as Facebook, Twitter, Mixi, etc., increasingly also among university students in Japan (Russel, 2012), numerous studies have emerged, which point out the benefits of utilising such networks for L2 learning (Lomicka & Lord, 2009; Stevenson & Lui, 2010; Wang & Vasquèz, 2012; Kent & Leaver, 2014). Blattner & Lomicka (2012) attested higher levels of motivation, affective learning, and a positive classroom climate. Kok (2008) and Rovai (2002) found that the use of SNS increases willingness to share data, information, and ideas, thus encouraging collaborative learning. In pedagogical contexts, however, it has to be ensured that SNS are used not only legally and ethically, but also in socially and culturally appropriate ways (Prichard, 2013).

Blattner & Fiori (2009) discovered that the use of SNS may also enhance learners’ socio-pragmatic skills, because they require the performance of speech acts such as greetings, thanks, requests, apologies, etc. In order to develop pragmatic competence, learners need both opportunities to practice communication as well as activities that raise their pragmatic awareness (Kasper & Rose, 2003). In this regard, Blattner & Fiori (2011) utilised the ‘Group’ application offered by Facebook, because it allows learners to observe authentic target language communication in other ‘open’ groups and provides them with the analytical tools necessary to inductively generalise about pragmatic aspects. Hanna & de Nooy (2003) observed that learners easily adapt to the conventions of communication regarding politeness, register, netiquette, and medium genre, and that they manage to interact in a productive manner with other individuals despite limited L2 abilities.

As far as the use of Facebook in L2 teaching in Japan is concerned, Promnitz-Hayashi (2013), who initiated group discussions by means of a private Facebook page, found that even more introverted students became actively engaged in tasks, and that the use of Facebook contributed in developing learner autonomy. In Dizon’s (2015) study, students appreciated the ease of use, convenience, and low-stress environment of using Facebook’s Group function for discussing pictures, videos, and links they shared with their classmates. In his action research study on using Facebook to provide learners with more opportunities to interact in their L2 English, Prichard (2013) found that Facebook is an effective tool for learners to develop digital literacy and to enhance L2 learning.

As far as German as a target language is concerned, Leier & Cunningham (2016) investigated differences between students’ private and teacher-assigned Facebook interactions at a university in New Zealand. As for GFL students in Japan, Waragai et al. (2014) found that learners make more use of the L2 within private messaging than
within assigned activities for concerns of formal adequacy. Clearly, more research is needed on both students’ private usage as well as study-related SNS use in order to arrive at an understanding of how learners might be assisted best in their endeavour to utilise such technologies for their L2 learning.

3. Participants and Procedures

The Facebook-project was carried out in a voluntary German class (CFER level A2) for third-year students with different majors consisting of two instruction units of 90 minutes per week. One unit was conducted in a CALL-classroom of which 45 minutes were devoted to the project. The participation in the project was voluntary and did not affect students’ grades. Nine of the eleven students participated in the project; two chose not to participate, but were assigned alternative tasks to be submitted by email instead.

In order to maintain students’ privacy and safety, a private Facebook Group page was set up by the teacher. This page was expected to serve as a platform for students to observe authentic target language communication in other ‘open’ groups as well as to engage in various speech acts themselves while sharing personal experiences with their classmates. Apart from the weekly assigned tasks, the students were also encouraged to use the Group page as a platform for class-related communication, such as apologising for a missed class, requesting or giving information or advice concerning their studies, as well as for other questions, announcements, or materials they wanted to share with this group.

At the beginning of the course each student received a speech act manual, which contained basic speech acts commonly used in online interaction (including greetings, thanks, requests, apologies, compliments, (well-)wishes, as well as phrases for expressing opinions, feelings, empathy, (dis-)like, (dis-)agreement, and surprise). The expressions compiled in the manual stem from GFL textbooks (CFER levels A1 to B1) as well as from the students’ own repertoire, which was established by a pre-course survey. Apart from serving as a source of reference during task performance, the speech act manual also contained space for students to write down examples or other relevant information concerning the use of the speech acts they encountered throughout the course.

The Facebook tasks involved accounts of the students’ experiences, feelings, or activities and aimed at eliciting different speech acts listed in Table 1. To facilitate the fulfilment of the tasks, pre-task activities were carried out in class to ensure that students were familiar with the relevant L2 expressions to perform the speech act(s) aimed at in the tasks. In order to encourage students to write more, the teacher sporadically also engaged in the Group’s timeline. In the lesson following the task completion, the students were asked to read their posts and comment on the timeline aloud, and the teacher provided corrective feedback placing particular emphasis on the correctness, adequacy, and appropriateness of the speech acts produced.

4. Task Aims and Analysis

The aim of the data analysis was twofold: 1) to determine what difficulties GFL learners face when producing certain speech acts in online communication and 2) to generate ideas for improving the tasks and the approach, so that students can be better assisted in this pursuit. The data collected for this study consist of the students’ posts and comments on the Group’s timeline as well as a pre- and post-course survey to get an insight into the students’ expectations and experiences concerning the project. Table 1 lists the ten tasks given throughout the course, the speech acts to be elicited in these tasks, and the criteria used for the analysis.
Table 1. Analysis of speech acts produced.

<table>
<thead>
<tr>
<th>Task</th>
<th>Speech acts to be elicited</th>
<th>Analysis</th>
</tr>
</thead>
<tbody>
<tr>
<td>1  Speech Act Search</td>
<td>any</td>
<td>none</td>
</tr>
<tr>
<td>2  Suggestions for a ‘Group’ photo</td>
<td>suggestions, (dis-)agreement</td>
<td>frequency grammaratical accuracy idiomatic appropriateness pragmatic appropriateness</td>
</tr>
<tr>
<td>3  Reporting on the spring vacation</td>
<td>(dis-)likes</td>
<td></td>
</tr>
<tr>
<td>4  Expressing feelings</td>
<td>feelings, (dis-)likes</td>
<td></td>
</tr>
<tr>
<td>5  Expressing one’s opinion on an article</td>
<td>opinions, (dis-)agreement</td>
<td></td>
</tr>
<tr>
<td>6  Describing activities</td>
<td>(dis-)likes</td>
<td></td>
</tr>
<tr>
<td>7  Expressing wishes</td>
<td>wishes, advices</td>
<td></td>
</tr>
<tr>
<td>8  Reporting on an experience</td>
<td>(dis-)likes</td>
<td></td>
</tr>
<tr>
<td>9  Requesting advice on learning German</td>
<td>requests, advices</td>
<td></td>
</tr>
<tr>
<td>10 Describing plans for the vacation</td>
<td>(well-)wishes</td>
<td></td>
</tr>
</tbody>
</table>

Task 1 differed from the other tasks inasmuch as it only addressed students’ passive pragmatic knowledge. In order to raise their awareness of L2 speech acts, they were exposed to authentic L2 online communication. For that purpose, they were asked to observe the interactions within an open Facebook-Group of their choice, to take screen shots of speech acts they identified, and to paste them on the timeline of the class’ Group page for discussion in class.

Tasks 2 to 10 required students to make more active contributions by writing their own posts according to the tasks given and by commenting on their classmates’ posts. As can be seen in Table 1, each of these tasks aimed at eliciting certain speech acts. In order to evaluate how adequate the tasks developed are regarding the elicitation of speech acts, the frequency of their occurrence per task was calculated. To gain further insights into the pragmatic competence of the learners, the speech acts they produced in their posts and comments were analysed according to grammatical and idiomatic correctness, as well as pragmatic appropriateness.

In order to determine students’ attitudes towards using Facebook for learning German and to ascertain to what extent they could improve their pragmatic skills throughout the project, a pre- and post-course survey was carried out. The survey contained open questions to generate ideas for improving the project as well as closed questions aimed at assessing students’ enjoyment and difficulties concerning task performance and determining the frequency of their SNS use by means of Likert scales.

5. Results

5.1. Students’ use of Facebook before the project

As far as the students’ use of Facebook before the project is concerned, the results of the pre-course survey showed that only five of the nine students who participated in the project had used Facebook before. They described themselves as rather ‘passive’ Facebook users. Judging from their ratings on a 4-point frequency scale (0, 1, 2, 3), their Facebook activities mostly consisted of checking message updates (1.8), reading the newsfeed (1.6), and liking friends’ posts (1.4); while functions that require more initiative or personal input received considerably lower averages, such as commenting on friends’ posts (0.6), chatting with friends (0.6), sharing information (0.6), sending messages (0.4), and writing one’s own posts (0.2). The survey also revealed that
students had no previous experience with using any kind of SNS for L2 learning, but most of them were optimistic regarding the Facebook-project as the following comments show: "SNS are helpful, because the communication with native speakers may accelerate our learning" and "SNS are convenient, because we have access to information, however, we need to use social media in an appropriate way."

5.2. Students’ task performance during the project

In order to assess the appropriateness of the tasks developed for this project, Table 2 shows the number of comments and speech acts (S.A.) elicited within the posts and comments of each task, as well as students’ perception of difficulty and enjoyment of the individual tasks determined by the post-course survey. For the enjoyment and difficulties the table displays average results obtained from ratings on a five-point frequency scale (-2, -1, 0, 1, 2).

Table 2. Quantitative findings concerning task performance.

<table>
<thead>
<tr>
<th>Task</th>
<th>Comments</th>
<th>S. A.</th>
<th>Difficulty</th>
<th>Enjoyment</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 Speech act search</td>
<td>---</td>
<td>---</td>
<td>-0.1</td>
<td>0.3</td>
</tr>
<tr>
<td>2 Suggestions for a ‘Group’ page photo</td>
<td>43</td>
<td>34</td>
<td>1.7</td>
<td>1.4</td>
</tr>
<tr>
<td>3 Reporting on the spring vacation</td>
<td>38</td>
<td>28</td>
<td>0.7</td>
<td>1.7</td>
</tr>
<tr>
<td>4 Expressing feelings</td>
<td>47</td>
<td>35</td>
<td>0.4</td>
<td>1.0</td>
</tr>
<tr>
<td>5 Expressing one’s opinion on an article</td>
<td>33</td>
<td>48</td>
<td>-1.2</td>
<td>1.0</td>
</tr>
<tr>
<td>6 Describing activities</td>
<td>57</td>
<td>48</td>
<td>0.3</td>
<td>0.9</td>
</tr>
<tr>
<td>7 Expressing wishes</td>
<td>70</td>
<td>71</td>
<td>0.0</td>
<td>1.2</td>
</tr>
<tr>
<td>8 Reporting on an experience</td>
<td>42</td>
<td>42</td>
<td>0.0</td>
<td>1.0</td>
</tr>
<tr>
<td>9 Requesting advice on learning German</td>
<td>19</td>
<td>33</td>
<td>-0.2</td>
<td>1.3</td>
</tr>
<tr>
<td>10 Describing plans for the vacation</td>
<td>71</td>
<td>58</td>
<td>1.4</td>
<td>1.7</td>
</tr>
<tr>
<td><strong>Averages</strong></td>
<td><strong>47</strong></td>
<td><strong>44</strong></td>
<td><strong>0.3</strong></td>
<td><strong>1.2</strong></td>
</tr>
</tbody>
</table>

As Table 2 shows, an average of 47 comments and 44 speech acts were produced per task; divided by the total number of nine participants this meant that on average each task prompted five speech acts and comments per student. The tasks ‘expressing wishes’ and ‘describing plans for the vacation’ generated most comments and speech acts and can therefore be regarded as appropriate for practicing pragmatics. As the figures for the individual tasks also indicate, the tasks of ‘requesting advice on learning German’, ‘expressing wishes’, and ‘expressing one’s opinion on an article’ elicited more speech acts than comments. Therefore, such tasks seem to be particularly appropriate for challenging the pragmatic competence of the learners. Relatively few speech acts compared to the number of comments were found in ‘reporting on the spring vacation’, ‘expressing feelings’, ‘describing activities’, and ‘describing plans for the vacation’.

As far as the difficulty of the individual tasks is concerned, there are considerable differences between the tasks. While ‘making suggestions for a Group photo’ and ‘describing plans for the vacation’ were perceived as rather easy, ‘requesting advice on learning German’ as well as the ‘speech act search’ were considered to be quite difficult. As for the enjoyment of task fulfilment, an average of 1.2 (based on ratings on a -2/2 scale) indicates that on average the tasks were seen as rather enjoyable. The notable
exception is the ‘speech act search’, which was rated as least enjoyable. Comparatively high scores for enjoyment were attained by the tasks of ‘describing plans for the vacation’ and ‘reporting on the spring vacation’.

To provide further insights into how accurately the speech acts were performed, Table 3 lists the types of speech acts produced by the students on the Group’s timeline according to the frequency of their occurrence (Total) and whether they were produced correctly or contained grammatical, idiomatic, or pragmatic inconsistencies.

**Table 3. Frequency and accuracy of speech acts performed.**

<table>
<thead>
<tr>
<th>Speech Act</th>
<th>Correct</th>
<th>Grammatically incorrect</th>
<th>Idiomatically inappropriate</th>
<th>Pragmatically inappropriate</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Expressing like</td>
<td>80</td>
<td>11</td>
<td>4</td>
<td>0</td>
<td>95</td>
</tr>
<tr>
<td>Advices</td>
<td>37</td>
<td>23</td>
<td>0</td>
<td>3</td>
<td>63</td>
</tr>
<tr>
<td>Agreements</td>
<td>41</td>
<td>5</td>
<td>3</td>
<td>3</td>
<td>52</td>
</tr>
<tr>
<td>Well-wishes</td>
<td>46</td>
<td>2</td>
<td>1</td>
<td>0</td>
<td>49</td>
</tr>
<tr>
<td>Expressing empathy</td>
<td>25</td>
<td>3</td>
<td>3</td>
<td>3</td>
<td>34</td>
</tr>
<tr>
<td>Expressing feelings</td>
<td>17</td>
<td>3</td>
<td>3</td>
<td>1</td>
<td>24</td>
</tr>
<tr>
<td>Thanks</td>
<td>19</td>
<td>2</td>
<td>2</td>
<td>0</td>
<td>23</td>
</tr>
<tr>
<td>Wishes</td>
<td>19</td>
<td>2</td>
<td>2</td>
<td>0</td>
<td>23</td>
</tr>
<tr>
<td>Requests</td>
<td>9</td>
<td>2</td>
<td>0</td>
<td>0</td>
<td>11</td>
</tr>
<tr>
<td>Opinions</td>
<td>7</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>8</td>
</tr>
<tr>
<td>Greetings</td>
<td>7</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>7</td>
</tr>
<tr>
<td>Expressing dislike</td>
<td>2</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>2</td>
</tr>
<tr>
<td>Compliments</td>
<td>2</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>2</td>
</tr>
<tr>
<td>Disagreement</td>
<td>2</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>2</td>
</tr>
<tr>
<td>Expressing surprise</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>Apologies</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td>315</td>
<td>54</td>
<td>18</td>
<td>10</td>
<td>397</td>
</tr>
<tr>
<td><strong>Percentage</strong></td>
<td>79%</td>
<td>14%</td>
<td>5%</td>
<td>3%</td>
<td>100%</td>
</tr>
</tbody>
</table>

Among the total of almost 400 speech acts produced during the tasks, likes, advice, well-wishes as well as expressions of empathy, and agreement appeared most frequently, while apologies, compliments, and expressions of dislike, disagreement, and surprises were used only rarely. Most of the speech acts were performed accurately. However, expressing feelings, empathy, and likes as well as giving advice or thanks, and making wishes and requests often entailed grammatical or idiomatic inadequacies. Pragmatic appropriateness seemed to be problematic for suggestions, advice, and agreements as well as for expressing feelings and empathy.
Considering the fact that the students had no previous experience using the L2 in social networking, it may come as a surprise that almost 80% of the speech acts on the timeline were performed accurately. This result may be partly attributed to the pre-task activities carried out in class, in which students could already practice the speech acts for the upcoming task. Some students also tried to 'play it safe' by only using tokens of speech acts they were already familiar with or by using the same expressions as their classmates; at times, however, including (spelling) mistakes.

Most inconsistencies in the speech acts were grammatical errors (14%), for example wrong verb conjugation in the compliment "Du fotografier[s]t gut! (You take good pictures!)") (Task 2), a missing particle in the advice "Ich empfehle dir, Aktien an[z]ukaufen und [zu] verkaufen! (I suggest to buy and to sell stocks!)") (Task 7), a missing object in the wish "Ich möchte [die Stadt] wieder besuchen! (I'd like to visit that city again!)") (Task 2), or a missing article in the agreement "Das ist [eine] gute Idee! (That's a good idea!)") (Task 9). Idiomatic mistakes, which only appeared in 5% of the speech acts, were for example the following expression of like, for which a wrong adjective was chosen, "Das ist fröhlich! [lustig]! (That looks happy [nice]!)") (Task 6) or the choice of a wrong verb in the agreement "Du bist [hast] recht! (You have [are] right!)") (Task 5).

Pragmatic inconsistencies, which only appeared in 3% of the data, are, for example, the choice of a wrong speech act as in the following reaction to a wish for good health "Bitteschön! (You're welcome!)") (Task 5) in which an expression of gratitude was mixed up with its acknowledgement. In other cases, it was not the type of speech act, which was wrong, but the token chosen for its performance. For example, in order to show 'genuine' compassion regarding someone's misfortune, the token "Schade! (Too bad!)") (Task 8) as a reaction to a post concerning rude treatment would be considered as too weak. In this case an expression which shows greater concern, such as "Das tut mir Leid für dich! (I feel sorry for you!)") would have been more appropriate. Similarly, the expression "Ich beneide dich! (I envy you!)") (Task 6) would be considered as too strong for commenting on a post containing a picture of a fruit shake. Here, less intrusive expressions, such as "Das sieht lecker aus! (That looks delicious!)") or "Du hast es gut (Lucky you!)") would be more common in German.

Finally, the choice of address forms may also contradict pragmatic norms. In a voluntary post, a student addressed me as "Herr Axel! (Mr. Axel!)") instead of "Lieber Axel (Dear Axel)". It is important to make students aware of how to use such address forms appropriately, because in online communication, where interlocutors do not necessarily know each other personally, they may cause offence. In this regard, the post-task activities conducted in this project proved to be quite effective, because they enticed learners to investigate cultural differences and language-specific nuances between the L2 items taught and related L1 expressions.

5.3. Project evaluation

As the quantitative analysis of the post-course student survey revealed, the implementation of the Facebook project as well as the usefulness of SNS as a tool for language learning were assessed positively (both with an average of 1.3 on a -2/2 scale). In their written comments, the students expressed an increase of their L2 skills regarding their pragmatic competence "I enjoyed the project, because I learnt new things in an entertaining way", "I learnt to express myself in a way that my classmates can understand" and "I learnt common German expressions, which hardly ever appear in textbooks or in the classroom". However, when performing the tasks, students strongly relied upon the expressions provided. Since most of the items taught were new to the students, some wished to have more time to explore nuances between them before actually using them. Other students pointed out social advantages of using Facebook "Since I normally don't talk to my teacher or classmates outside the class, it was a good chance for us to get to know each other better." and "It was interesting to
find out what my classmates are doing in their free time, and what they think on certain subjects.”

However, some critical comments also emerged, for example concerning the use of the medium Facebook itself “Since I normally don’t use Facebook, I found it hard to get into the habit of checking the newsfeed regularly and to comment on my classmates’ posts.” For others, performing the tasks posed problems because they either lacked the required L2 skills to fulfil the task as in “I wanted to comment on my classmates’ posts, but I did not know how to express myself in German” or they encountered cultural difficulties “I found it difficult to express my feelings and opinions”.

6. Summary and Discussion

The results of this explorative study will be summarised according to the three research questions listed in section 1. As far as students’ response to using Facebook for L2 learning is concerned (research question 1), the learners in this study seem to have enjoyed the project as their resourceful posts and comments on the Groups’ timeline, their active participation in pre- and post-task activities, and their overwhelmingly positive comments in the post-course survey show. Taking into account that learners in Japan are rather shy, the use of Facebook encouraged them to apply their L2 knowledge and to state their opinions more freely than they would normally do in face-to-face interactions, which has also been observed in the studies conducted by Promnitz-Hayashi (2013) and Dizon (2014). They felt encouraged to share personal information with their classmates and enjoyed completing the tasks collaboratively, which increased their motivation and led to a positive classroom climate, confirming observations made by Kok (2008) and Blattner & Lomicka (2012). As the comparison of students’ comments in the pre- and post-course survey also revealed, those students who were already frequent users of SNS, profited most from this blended learning approach. As pointed out by Prichard (2013), learners of the net generation may have little technical difficulty getting accustomed to the use and functionality of SNS, but they need guidance on how to use them in culturally appropriate ways.

Regarding the appropriateness of the approach used in this project (research question 2), the tasks developed seem to have been effective for eliciting a large number of different speech acts. Judging from the students’ feedback, most tasks were rated as ‘enjoyable’ since they related to their interests and activities. However, it has to be admitted that the students’ posts did not purely derive from an authentic desire to socialise or to share information, but were a requirement of their language class. In this regard, it has to be further investigated how differences between students’ private and teacher-assigned SNS interactions affect L2 learning (Leier & Cunningham, 2016; Waragai et al., 2014). In their more learner-centred approach, Promnitz-Hayashi (2013) let students design their own tasks, which significantly increased the number of prompted comments as compared to teacher-assigned tasks. As for the speech act search (Task 1), the learners in this study perceived it as rather difficult and far less enjoyable than the other tasks. It was hoped that this task would entice students to incorporate speech acts they encountered in authentic L2 interactions into their own repertoire. However, as their contributions on the Group’s timeline revealed, they mostly relied on the expressions provided in pre-task activities.

As far as the potential offered by Facebook for acquiring L2 pragmatics (research question 3) is concerned, the overall results of this case study are in line with Blattner & Fiori’s (2009 and 2011) findings, which highlight the benefits of observation-based awareness-raising tasks for the development of pragmatic competence. While the tasks developed for this study placed a stronger focus on speech act production, rather than observation, the approach chosen proved to be equally beneficial for improving learners’ pragmatic skills. As also observed by Hanna & de Nooy (2003), the learners in this study managed to interact effectively with each other despite their limited L2 abilities. However, to what extent they will be able to communicate in a pragmatically appropriate manner outside the ‘guided’ Group framework of this blended learning
approach, still remains to be determined. In order to get a deeper insight into learners’
acquisition and development of pragmatic skills, finer research tools are required, such
as introspective interviews with learners, think-aloud protocols of their writing
processes, as well as longitudinal and comparative studies. Most of the studies on the
use of SNS for language learning to date, including the project described in this paper,
are exploratory in nature. More empirical research is needed to confirm the conclusion
of these studies.

References

Possibilities. *Instructional Technology and Distance Learning (ITDL)*, 6(1), 17-28.

language learners’ development of socio-pragmatic awareness and multiliteracy skills.
*CALICO Journal*, 29(1), 24-43.

Blattner, G. & Lomicka, L. (2012). Facebook-ing and the social generation: A new era of


Hanna B.E. & de Nooy, J. (2003). A funny thing happened on the way to the forum:
Electronic discussion and foreign language learning. *Language Learning & Technology,
7*(1), 71-85.


Watts, T. Muller, & M. Swanson (Eds.), *JALT2007 Conference Proceedings. Tokyo: JALT,
1023-1032.*

Blackwell.

Instructional Technology and Distance Learning, 5(10), 25-32.

Kent, M. & Leaver, T. (2014). *An Education in Facebook?: Higher Education and the

Leier, V. & Cunningham, U. (2016). "Just facebook me": A study on the integration of
Facebook into a German language curriculum. *CALL communities and culture – short
papers from the EUROCALL 2016 Conference held in Limassol, Cyprus*, pp. 260-264.
Research-publishing.net. doi:10.14705/rpnet.2016.eurocall2016.572.

Collaboration in Foreign Language Learning.* San Marcos, Texas: CALICO.


*CALICO Journal, 30*(2), 204-225.

Cambridge University Press.

Rovai, A.P. (2002). Sense of community, perceived cognitive learning, and persistence
in asynchronous learning networks. *Internet and Higher Education, 5*, 319-332. doi:
10.1016/S1096-7516(02)00130-6.


Testing audiovisual comprehension tasks with questions embedded in videos as subtitles: a pilot multimethod study

Juan Carlos Casañ Núñez
Universitat Politècnica de València, Spain

Abstract

Listening, watching, reading and writing simultaneously in a foreign language is very complex. This paper is part of wider research which explores the use of audiovisual comprehension questions imprinted in the video image in the form of subtitles and synchronized with the relevant fragments for the purpose of language learning and testing. Compared to viewings where the comprehension activity is available only on paper, this innovative methodology may provide some benefits. Among them, it could reduce the conflict in visual attention between watching the video and completing the task, by spatially and temporally approximating the questions and the relevant fragments. The technique is seen as especially beneficial for students with a low proficiency language level.

The main objectives of this study were to investigate if embedded questions had an impact on SFL students' audiovisual comprehension test performance and to find out what examinees thought about them. A multimethod design (Morse, 2003) involving the sequential collection of three quantitative datasets was employed. A total of 41 learners of Spanish as a foreign language (SFL) participated in the study (22 in the control group and 19 in the experimental one). Informants were selected by non-probabilistic sampling. The results showed that imprinted questions did not have any effect on test performance. Test-takers' attitudes towards this methodology were positive. Globally, students in the experimental group agreed that the embedded questions helped them to complete the tasks. Furthermore, most of them were in favour of having the questions imprinted in the video in the audiovisual comprehension test of the final exam. These opinions are in line with those obtained in previous studies that looked into experts', SFL students' and SFL teachers' views about this methodology (Casañ Núñez, 2015a, 2016a, in press-b). On the whole, these studies suggest that this technique has potential benefits for FL learning and testing. Finally, the limitations of the study are discussed and some directions for future research are proposed.

Keywords: Audiovisual comprehension, listening comprehension, multimethod design, Spanish as a foreign language, subtitles, video listening test.
in which listeners select and interpret information which comes from auditory and visual cues in order to define what is going on and what the speakers are trying to express”. Thus, it is considered that factors such as “proxemics, kinesics and deictics are all part of the message. They are not just a sort of gloss on the verbal component” (Riley, 1979, p. 84). Harris (2003), Lynch (2012) and Riley (1979) have suggested that the term listening does not reflect the multimodal nature of most listening comprehension situations. From now on, in order to show the dual dimension of this communicative activity, the compound listening/audiovisual comprehension will be employed.

According to a number of authors (Lynch, 2009; Mendelsohn, 1994; Rubin, 1995a, Ur, 1999), video materials should prevail over audio recordings to practice listening/audiovisual comprehension. To begin with, it is consistent with a definition of the skill as a process of interpretation of auditory and visual information. Besides, it allows the learner to observe the reality of most speaking interactions. In addition to that, video has a positive effect on motivation (Flowerdew & Miller, 2005; Ur, 1994, 1999; Vandergrift & Goh, 2012). Finally, there are arguments in favour of multimodal learning. According to the cognitive theory of multimedia learning (CTML), multimedia "takes advantage of the full capacity of humans for processing information. When we present material only in the verbal mode, we are ignoring the potential contribution of our capacity to also process material in the visual mode" (Mayer, 2014, p. 6). Thus, as stated by the multimedia principle of the CTML "students learn better from words and pictures than from words alone" (Mayer, 2001, p. 63). This theory was not developed specifically for learning foreign languages, however, some principles are applicable to this field (Plass & Jones, 2005). In relation to the multimedia principle, these authors point out that "it is the combination of both visual and verbal presentations of information that has most strongly and consistently supported listening and reading comprehension and vocabulary acquisition" (p. 479). Video materials should also prevail over audio recordings for testing listening/audiovisual comprehension. First, it is congruent with the double nature of this communicative activity. Second, it is in harmony with common practice in the classroom (Buck, 2001; Gruba, 1997; Pardo-Ballester, 2016). Third, it increases the validity of the test (Bejar, Douglas, Jamieson, Nissan, & Turner, 2000; Wagner, 2007, 2008, 2010a), its authenticity (Alderson, 2005; Bejar et al. 2000; Ockey, 2007; Wagner, 2007, 2008) and its naturalness (Alderson, 2005). Fourth, “seeing the situation and the participants tends to call up relevant schemes” (Buck, 2001, p. 172). Lastly, not using video in language testing may have a negative backwash effect. If the skill is tested only using audio recordings, then there is a pressure to practice this communicative activity mainly with this sort of materials.

When designing listening/audiovisual comprehension tasks, it is essential to take into account that “it is extremely difficult to listen and write at the same time, particularly in a foreign language” (Underwood, 1989, p. 48), and that listening, viewing, reading and writing at the same time can be even more difficult. On the one hand, there is a conflict of visual attention between viewing a video and completing a written activity at the same time. On the other hand, it should be borne in mind that “there is universal agreement that working memory when dealing with novel information is very limited in capacity” (Sweller, Ayres & Kalyuga, 2011, p. 42). As Vandergrift and Goh (2012) highlight, paying attention to the video and the task simultaneously may cause working memory overload. This complexity helps to explain the relatively low degree of attention paid to the video in studies that researched test-takers’ viewing rates during foreign language (FL) video-based listening comprehension tests. Ockey (2007) does not supply the average watching rate but, from the data, it can be calculated that it is 44.9%. Wagner (2007) found that examinees made eye contact with the screen 69% of the time, while in a later study (Wagner, 2010b), this figure decreased to 47.9%. These authors recorded the participants while taking the tests and they measured the amount of time informants look towards the monitor. Suvorov (2015) uses eye-tracking technology to investigate how examinees interact with two types of videos: context videos and content videos. He discovers that test-takers spend 58% of the time
watching content videos and 51% context videos. These low degrees of attention to the image may have a negative influence both on understanding the video and on the development of the communicative activity. The complexity of the while-viewing phase also contributes to explaining the mixed conclusions in (a) studies that compared the results obtained by a group after conducting a video-based FL test with those achieved by another group that took an audio-only version of the same test, and (b) research into test-taker’s attitudes towards the image. Some authors do not find significant differences in performance between taking a test with video and completing the same test with an audio-only version of the audiovisual text (Batty, 2014; Coniam, 2001; Gruba, 1993; Londe, 2009), other authors discover that test-takers achieve higher scores with the video based test (Sueyoshi & Hardison, 2005; Wagner, 2010a, 2013), and Suvorov (2008) finds evidence that results are higher with the audio-only version. Similarly, in some studies, examinees have a positive attitude towards the video image (Sueyoshi & Hardison, 2005; Wagner, 2010b), and in others they have a negative attitude toward the video (Alderson, Clapham & Wall, 1995; Coniam, 2001; Suvorov, 2008). Of course, other elements may have played a role in these mixed results. These include the focus of listening/audiovisual questions, the complexity of the task and the text, the way in which the viewing was carried out, the quality of the input, the stress generated by the tests, the greater or lesser degree of solidarity between visual and verbal information, the type of visuals (content or context visuals), the influence of the video cameras on viewing behaviour, and so on.

In order to keep while-viewing work manageable, it is advisable that written tasks involve scarce reading and writing, and that they require few active elements to be stored in working memory. It is also recommendable to use the technique of paused listening/viewing (see Field, 2008; Stempleski & Tomalin, 2001; Stoller, 1992). Thirdly, it is useful to spatially approximate the video and the activity to reduce the time needed to shift from one stimulus to another (from the task to the video and vice versa). Lastly, as proposed in this study, it may be more beneficial for the learner to see the comprehension questions embedded in the video in the form of subtitles and synchronized with the relevant fragments (Casañ Núñez, 2015a). Basically, comprehension questions appear on screen a few seconds before the beginning of the fragment to which they are related, they remain visible for the duration of the relevant snippet and they disappear when the relevant part of the video finishes. Compared to viewings where the activity is available only on paper, this technique could minimize conflict in visual attention by spatially and temporally approximating the questions and the pertinent scenes, and it could reduce the cognitive strain of the task, since students would only need to pay attention to one subtitled question at a time, instead of different printed questions on paper. According to Field (2008), unskilled listeners may reach working memory overload faster than competent listeners because the former have poorly automatized decoding processes and they spend a great deal of working memory on decoding. Consequently, the procedure is seen as especially beneficial for students with a low proficiency language level. In addition, it can be used occasionally in higher proficiency levels for two main reasons: firstly, it helps learners keep focused on what they are watching compared to viewings with questions on paper; secondly, the study reported in Casañ Núñez (2016a) shows that learners with a high listening/audiovisual comprehension level in SFL (approximately B2+/C1) have positive views on this technique.

This procedure may be employed in different contexts. It is suitable for paper-and-pencil listening/audiovisual comprehension tasks for language learning or testing (see Figure 1) and for CALL or CALT (see Figures 2 and 3). As for learning and testing activities, Alderson et al. (1995, p. 42) state that “the main difference between a test and an exercise is that with exercises learners get support: with tests, they do not”. A video demonstration of this technique is available from https://youtu.be/ALw8XJkrbDQ (01/02/2017).
Figure 1. Example of an audiovisual comprehension question embedded in the video in the form of a subtitle. "¿De qué temas habla el chico pelirrojo?" [What is the redheaded boy talking about?] From the Spanish film Los peores años de nuestra vida by Emilio Martínez Lázaro.

Figure 2. Prototype of an audiovisual comprehension task for CALL. Notice that playback controls are available to learners. "¿Cómo se llama la chica?" [What is the girl's name?]. From the Spanish film Ópera prima by Fernando Trueba.
Previously, a theoretical framework describing the use of this technique and its potential benefits and limitations has been proposed (Casañ Núñez, 2015a). The framework was developed out of a literature review, the teaching experience with this procedure and the comments of a group of experts in teaching Spanish as a Foreign Language (SFL), Spanish linguistics and/or the use of technology. Also, small studies have been carried out to investigate what SFL university teachers and university students with a high listening/audiovisual comprehension level in SFL (approximately B2+/C1) think about this technique (Casañ Núñez, 2016a, in press-b). The results suggest that, overall, teachers and learners have positive views about this methodology. In addition to experts’, teachers’ and students’ views, it is fundamental to find out what effect this technique has on learners’ audiovisual comprehension and viewing behaviour.

The main purposes of the current pilot study were to investigate if the technique had an impact on SFL students’ audiovisual comprehension test performance and to find out what examinees thought about imprinted questions. Moreover, it explored some learner preferences regarding listening/audiovisual comprehension. The study used datasets from previous research that described the development of a listening/audiovisual test (Casañ Núñez, 2016b, pp. 36-51). The study reported in this paper, however, had different objectives; it took into account data that was not inspected and it analysed the data addressing the following research questions:

Research question 1: How important is it for learners to practise listening/audiovisual comprehension in the classroom? What type of recordings (audio or video) do students prefer for practising listening/audiovisual comprehension in the classroom? Do learners think that the visual input helps them to understand what speakers are saying? How do students practise listening/audiovisual comprehension outside the classroom?

Research question 2: Does the use of questions embedded within the video in the form of subtitles and synchronized with the relevant fragments in a FL audiovisual comprehension test facilitate test-takers’ performance? Do the test-takers of the experimental group score higher or lower than the test-takers of the control group?
It was hypothesized that examinees that take the audiovisual comprehension test with questions embedded in the video, that is, students in the experimental group, would outperform examinees that took the same test without them, i.e., those in the control group. As described in the introduction, imprinted questions could minimize the conflict in visual attention between watching the video and completing the task, by spatially and temporally approximating the questions and the relevant fragments. Moreover, this technique could diminish the cognitive strain of the activity, because learners would only need to pay attention to one request each time, instead of several printed questions. Nonetheless, a null hypothesis of no difference was tested.

Research question 3: What are test-takers’ attitudes towards the use of questions embedded in the video as subtitles?

- Research question 3.1: To what extent do test-takers in the experimental group agree or disagree that embedded questions helped them to complete the tasks?
- Research question 3.2: Is there any subtitled question consistently considered more or less helpful than the others by the experimental group?
- Research question 3.3: Are students in the experimental group in favour of having the questions embedded in the video (in addition to having them on paper) in the audiovisual comprehension test of the final exam?

As mentioned in the introduction, the use of subtitled questions has potential benefits. In addition, a previous study (Casañ Núñez, 2016a) showed that SFL learners had positive attitudes towards this methodology. Therefore, in this study, it was hypothesized that students in the experimental group would agree that the questions embedded in the form of subtitles aided them, and that they would be in favour of having them in the test of the final exam.

2. Method

2.1. Study design

A multimethod design (Morse, 2003) was employed. It involved the sequential collection of three quantitative datasets that were used basically to answer different subquestions. First, participants were surveyed with the purpose of getting to know the sample and some of their preferences regarding listening/audiovisual comprehension. Second, an audiovisual comprehension test with two variants was administered to find out if there were differences in performance between test-takers of the control and experimental groups. Third, attitudinal data towards the use of questions embedded in the form of subtitles from the experimental group was collected through a questionnaire.

2.2. Participants

SFL students were selected by a convenience, non-probabilistic sampling method (Dörnyei, 2007, pp. 98-99). The instruments were administered in different lessons. Some students did not complete the first questionnaire because they missed the lessons where they were handed out (see Table 1).

Table 1. Number of informants that completed each instrument.

<table>
<thead>
<tr>
<th>Instruments</th>
<th>Groups</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Control</td>
</tr>
<tr>
<td>First questionnaire</td>
<td>18</td>
</tr>
<tr>
<td>Audiovisual comprehension test</td>
<td>22</td>
</tr>
<tr>
<td>Second questionnaire (only for the experimental group)</td>
<td>n/a</td>
</tr>
</tbody>
</table>
All participants were enrolled in Spanish II, a foreign language course delivered at the Universidade de Coimbra (Portugal). Spanish II had three shifts. Two of them had fewer persons registered. Thus, it was decided that the smaller groups completed the same version of the test. Randomly, the less numerous shifts were designated control groups and the remaining one, the treatment group. All participants had the same Spanish teacher and studied Human or Social Sciences degrees. All but two informants were between eighteen and twenty-four years old. All were Lusophones except for one participant in the experimental group who was of Ukrainian origin. Roughly speaking, students had been studying Spanish for a similar amount of time and they had spent an analogous amount of time in Spanish-speaking countries. 83.3% of the participants of the control group and 88.9% of the informants of the experimental group reported that they were very interested or very much interested in cinema. The others answered “neutral”. That aspect was relevant because the videotexts employed in the audiovisual comprehension test were film scenes and a low interest in cinema could have some negative influence on performance. A pre-test to determine the informants’ level of audiovisual comprehension in Spanish was not administered. From the observation of the texts and tasks employed for practising listening/audiovisual comprehension in the classroom, it was estimated that most students had a B1+ level in this skill (according to the Common European Framework for Languages). As described in Casañ Núñez (2016b, pp. 39-40), in order to estimate the degree of equivalence between the groups regarding audiovisual comprehension, participants completed slightly modified versions of the first and second tasks from a B1 level listening test designed by Hidalgo de la Torre (Coord., 2013: pp. 52-53). Both groups obtained similar high scores. The average mark in the control group was 9.566 out of 12 (SD = 1.861) and in the experimental one, it was 9.684 out of 12 (SD = 1.827). That suggested that there were few differences between the groups and that the B1 level listening tasks were easy for test-takers. The Mann-Whitney test for two independent samples confirmed that there were no statistically significant differences between the scores of the groups (U = 146.5, z = -.185, p = .853, r = -.031).

2.3. Instruments

To collect the data, three instruments were developed: two questionnaires and an audiovisual comprehension test with two versions. The first questionnaire aimed at gathering specific information about the sample group: sociodemographic aspects, what importance was attributed to the exercise of listening/audiovisual comprehension in the classroom to learn Spanish, some learning preferences about listening/audiovisual comprehension, to what extent video images were considered helpful or unhelpful to understand the interlocutors, and how listening/audiovisual comprehension was practiced outside the classroom. It included twenty-six items and, following the classification of Saris and Gallhofer (2014), it employed open requests and closed categorical requests. The development of the instrument involved expert review, a pilot, and a study where the repeated-surveys method (Brown, 2001, pp. 171-172) was used to calculate its reliability. The questionnaire and a detailed account of its elaboration can be found in Casañ Núñez (in press-a).

The audiovisual comprehension test had two variants: a traditional one for the control group and an experimental one for the treatment group. An extensive account of the planning, design and trialling of both versions of the test can be found in Casañ Núñez (2016b). In both instances, tasks were available on paper. Besides, in the experimental one, questions were embedded in the video in the form of subtitles and were synchronized with the relevant fragments (see Figures 4 and 5).
The target language use domain was the comprehension of informal conversations pertaining to the personal domain in Spanish romantic comedies. The construct measured three skills: extracting specific information, identifying general ideas and recognizing feelings in face-to-face informal conversations. The test was composed of two tasks, two texts and seven items. The main features of the texts and the experimental items can be seen in Tables 2 and 3. To indicate the film scene, an eight-digit time code is used. The first three pairs of digits correspond to hours, minutes and seconds, respectively, and the last pair, to frames. Thus, 00:02:36:12 designates the following point: 2 minutes, 36 seconds and 12 frames of Ópera prima.

**Table 2. Main characteristics of the input texts.**

| 1. Text source | Videotext (source: CEFR a p. 49) Scene from the comedy Ópera prima by Fernando Trueba where an informal conversation takes place. Location: from 00:00:00:00 to 00:02:36:12. | Videotext Scene from the comedy Los peores años de nuestra vida by Emilio Martínez Lázaro where an informal conversation takes place. Location: from 00:11:10:00 to 00:12:48:15. |
| 2. Authenticity | Genuine | Genuine |
| 3. Domain type (source: CEFR* page 45) | Personal | Personal |
| 4. Text length | 2 min and 36 s | 1 min and 38 s |
The EUROCALL Review, Volume 25, No. 1, March 2017

Table 3. Description of the embedded questions in the scenes from Ópera prima (1-4) and Los peores años de nuestra vida (4-7).

<table>
<thead>
<tr>
<th>Subtitle</th>
<th>Focus</th>
<th>Question type</th>
<th>Timing</th>
<th>No. of lines</th>
<th>Font and size</th>
<th>Colour</th>
<th>Estimated level</th>
</tr>
</thead>
<tbody>
<tr>
<td>¿Cómo se llama la chica?</td>
<td>ESI</td>
<td>Multiple-choice</td>
<td>00:01:10:00 00:01:37:09</td>
<td>1</td>
<td></td>
<td></td>
<td>A1</td>
</tr>
<tr>
<td>¿De qué hablan?</td>
<td>IGI</td>
<td>Multiple-choice</td>
<td>00:01:44:14 00:02:02:02</td>
<td>1</td>
<td></td>
<td></td>
<td>A2/B1</td>
</tr>
<tr>
<td>¿Cuál es el teléfono de la chica?</td>
<td>ESI</td>
<td>Open question with only one possible answer</td>
<td>00:02:04:04 00:02:22:00</td>
<td>1</td>
<td>Arial Narrow 36</td>
<td>White (hex colour code #FFFFFF)</td>
<td>A2</td>
</tr>
<tr>
<td>¿Qué sentimientos puede tener el chico por la chica?</td>
<td>RF</td>
<td>Open</td>
<td>00:02:27:08 00:02:36:12</td>
<td>2</td>
<td></td>
<td></td>
<td>A2/B1</td>
</tr>
<tr>
<td>¿De qué temas habla el chico pelirrojo?</td>
<td>IGI</td>
<td>Multiple-choice</td>
<td>00:11:40:00 00:12:25:24</td>
<td>1</td>
<td></td>
<td></td>
<td>B2</td>
</tr>
<tr>
<td>¿Cómo se llama la chica?</td>
<td>ESI</td>
<td>Multiple-choice</td>
<td>00:12:39:04 00:12:43:24</td>
<td>1</td>
<td></td>
<td></td>
<td>A1</td>
</tr>
<tr>
<td>¿Qué sentimientos puede tener el chico pelirrojo por la chica?</td>
<td>RF</td>
<td>Open</td>
<td>00:12:45:00 00:12:48:15</td>
<td>2</td>
<td></td>
<td></td>
<td>B2</td>
</tr>
</tbody>
</table>

Abbreviations: ESI - extracting specific information. IGI - identifying general ideas. RF - recognizing feelings.

The second questionnaire aimed at answering the third group of research questions. It was a development of a survey used in a previous exploratory study (Casañ Núñez, 2016a). A questionnaire design was chosen because they are "uniquely capable of gathering large amounts of information quickly" (Dörnyel, 2007, p. 101). The instrument had three parts: title, introductory text and items (see appendix). The title tried to be as informative as possible. The introductory text specified the objective of the questionnaire, it mentioned that there were no correct or incorrect answers, and it stated that the data collected would be treated confidentially and would only be used for academic purposes. Following the classification of Saris and Gallhofer (2014), the instrument was made up of two types of items: closed categorical requests and open
requests. The former ones were of two subtypes: requests with nominal response categories and requests with ordinal response categories.

As recommended by Saris and Gallhofer (2014), the response options of the closed categorical requests attempted to offer a range of options, broad enough to encompass the possible answers of participants, and they were mutually exclusive. Closed categorical requests with ordinal response categories consisted of Likert-type items (1, 2.1, 2.2, 2.3, 2.4, 5, 6.1, 6.2 and 6.3). They made up a scale that attempted to answer research question 3.1 *To what extent do test-takers in the experimental group agree or disagree that embedded questions helped them to complete the tasks?* The scale was named *helpfulness scale*. Additionally, individual items tried to respond to research question 3.2 *Is there any subtitled question ranked consistently higher or lower than the others by the experimental group?* Test-takers expressed their opinion about the tasks of the test in a global way (items 1 and 5) and about each subtitled question individually (items 2.1, 2.2, 2.3, 2.4, 6.1, 6.2 and 6.3). There are discrepancies on the number of anchors of Likert scales: Likert (1932) employs mainly five alternatives; Jacoby and Matell (1971) suggest that three categories are sufficient; Allen and Seaman (2007) advise a minimum of five options, and Bisquerra and Pérez-Escoda (2015) recommend eleven. Three anchors were opted for, primarily for two reasons. First, it was enough to find out if the attitude towards subtitled questions was positive, negative or neutral. Second, the size of the sample was small and it would be necessary to collapse the categories. The alternatives chosen were: agree, neither agree nor disagree and disagree. In that selection there was a balance between positive and negative alternatives and a neutral central point could be identified. In addition, terms that indicated total agreement or disagreement were avoided because, according to Cohen, Manion and Morrison (2011), people often prefer to skip them so as not to appear extremist. Closed categorical requests with nominal response categories (items 3 and 7) sought to answer research question 3.3 *Are students in the experimental group in favour of having the questions embedded in the video (in addition to having them on paper) in the audiovisual comprehension test of the final exam?*

"By permitting greater freedom of expression, open-format items can provide a far greater richness than fully quantitative data. The open responses... can also lead us to identify issues not previously anticipated“ (Dörnyei, 2007, p. 107). Their main drawback lies in the complexity of analysing them (Cohen et al., 2011). Two items of this nature were included so that participants could express their opinion freely (items 4 and 8). In order for test-takers’ production level in Spanish not to restrict responses, the possibility of writing both in Spanish and Portuguese was offered.

### 2.4. Procedures

All instruments were administered by the researcher during *Spanish II* lessons. The fact that the author was also the teacher may have favoured the fact that participants took the questionnaires and the test seriously (an essential requirement for the results to be valid). The first questionnaire was administered for the first time in February 2014 to students from the three shifts. Precise instructions were provided and students were told that there were no correct or incorrect replies. No time limit was imposed and questions arising from any doubts were answered. Informants took up to 8 minutes to complete the questionnaire and none of them requested help.

The test and the second questionnaire were administered in the same classroom in March 2014. Treatment groups completed the experimental version of the test and the questionnaire, whereas the control group took the traditional variant of the test. Test-takers were informed that they could take as long as they needed to answer. The slowest test-taker of the control group employed 12 minutes and 27 seconds to finish the test. The slowest examinees of the two experimental groups needed 20 minutes and 14 seconds, and 19 minutes and 42 seconds to respond to the test and the survey. A detailed account of the procedures followed can be found in Casañ Núñez (2016b, pp. 42-43).
2.5. Data analyses

All quantitative statistical analyses were carried out using SPSS version 21, except for effect sizes. For that purpose, the Windows scientific calculator was used. Data was double-checked for accuracy to avoid input errors. Furthermore, a frequency analysis of all variables was carried out to verify that there were no missing or anomalous values. Table 4 sums up the quantitative data analyses that were followed to answer each research question.

As for the open-ended items in the second questionnaire (4 and 8), only 5 out of 19 participants in the experimental group made comments, and those observations were short (one sentence). This might have to do with Dörnyei’s cautionary advice (2007, p. 105): questionnaires “are unlikely to yield the kind of rich and sensitive description of events and participant perspectives that qualitative interpretations are grounded in”. As the amount of qualitative data was so small, the only analysis consisted in quantizing related responses.

Table 4. Quantitative data analyses used for answering research questions

<table>
<thead>
<tr>
<th>Research question</th>
<th>Research instrument</th>
<th>Analyses</th>
</tr>
</thead>
<tbody>
<tr>
<td>RQ1</td>
<td>First questionnaire (instrument developed by Casañ Núñez, in press-a)</td>
<td>Frequencies</td>
</tr>
<tr>
<td>RQ2</td>
<td>Audiovisual comprehension test (instrument developed by Casañ Núñez, 2016b)</td>
<td>Scores of each version of the test: descriptive statistics, analyses of facility, discrimination and reliability, and correlations with two tasks from a listening exam (see Casañ Núñez, 2016b, pp. 44-51)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Tests scores: Shapiro–Wilk test</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Comparison of tests scores: Mann–Whitney test</td>
</tr>
<tr>
<td>RQ3</td>
<td>Second questionnaire (based on a survey used by Casañ Núñez, 2016a)</td>
<td>RQ 3.1: Reliability analysis of the helpfulness scale (Cronbach’s alpha, corrected item-total correlation, Cronbach’s alpha if item deleted) and descriptive statistics of the scale</td>
</tr>
<tr>
<td></td>
<td></td>
<td>RQ 3.2: Friedman test (items 2.1, 2.2, 2.3, 2.4, 6.1, 6.2 and 6.3) and two Wilcoxon Signed Rank tests using a Bonferroni adjusted alpha value (item 2.4 with item 2.2, item 2.4 with item 2.3).</td>
</tr>
<tr>
<td></td>
<td></td>
<td>RQ 3.3: frequencies (items 3 and 7)</td>
</tr>
</tbody>
</table>

3. Results and discussion

3.1. Research question 1

The first research question addressed learners’ thoughts and preferences regarding listening/audiovisual comprehension. Informants’ answers can be found in Tables 5 to 8.
Table 5. How important is it for learners to practise listening/audiovisual comprehension in the classroom?

<table>
<thead>
<tr>
<th></th>
<th>Frequency</th>
<th>Percent</th>
<th>Valid Percent</th>
</tr>
</thead>
<tbody>
<tr>
<td>Valid</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Very*</td>
<td>16</td>
<td>39.0</td>
<td>44.4</td>
</tr>
<tr>
<td>Very much</td>
<td>20</td>
<td>48.8</td>
<td>55.6</td>
</tr>
<tr>
<td>Total</td>
<td>36</td>
<td>87.8</td>
<td>100.0</td>
</tr>
<tr>
<td>Missing**</td>
<td>System</td>
<td>5</td>
<td>12.2</td>
</tr>
<tr>
<td>Total</td>
<td>41</td>
<td>100.0</td>
<td></td>
</tr>
</tbody>
</table>

*Participants could choose one of five responses: (a) very little, (b) a little, (c) neutral, (d) very, and (e) very much.

**Missing values correspond to informants that did not complete the first questionnaire.

All informants considered that practising listening/audiovisual comprehension in the classroom was either “very” important (16 / 44.4%) or “very much” important (20 / 55.6%). These results are congruent with the weight of listening/comprehension in communication (“we listen to twice as much language as we speak, four times as much as we read, and five times as much as we write”, Celce-Murcia & Olshtain, 2000, p. 102) and in acquisition (“in order for acquirers to progress to the next stage in the acquisition of the target language, they need to understand input language that includes a structure that is part of the next stage”, Krashen & Terrel, 1983, p. 32).

Table 6. What type of recordings (audio or video) students prefer for practising listening/audiovisual comprehension in the classroom?

<table>
<thead>
<tr>
<th></th>
<th>Frequency</th>
<th>Percent</th>
<th>Valid Percent</th>
</tr>
</thead>
<tbody>
<tr>
<td>Valid</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Audio*</td>
<td>4</td>
<td>9.8</td>
<td>11.1</td>
</tr>
<tr>
<td>Video</td>
<td>11</td>
<td>26.8</td>
<td>30.6</td>
</tr>
<tr>
<td>Both equally</td>
<td>21</td>
<td>51.2</td>
<td>58.3</td>
</tr>
<tr>
<td>Total</td>
<td>36</td>
<td>87.8</td>
<td>100.0</td>
</tr>
<tr>
<td>Missing**</td>
<td>System</td>
<td>5</td>
<td>12.2</td>
</tr>
<tr>
<td>Total</td>
<td>41</td>
<td>100.0</td>
<td></td>
</tr>
</tbody>
</table>

*Participants could choose one of three options: (a) audio, (b) video, and (c) both equally.

**Missing values correspond to informants that did not complete the first questionnaire.

Most students (21 / 58.3%) reported that they did not have a preference for audio or video materials for practising listening/audiovisual comprehension in the classroom. In addition, many more participants showed a predilection for video (11 / 30.6%) instead of audio (4 / 11.1%). As participants were not asked to justify their answer, it is not possible to know why they made their choices. Possibly, the preference for video over audio could be due to the positive effect of video on motivation (Flowerdew & Miller, 2005; Ur, 1994, 1999; Vandergrift & Goh, 2012). Despite this, from the point of view of the learners’ preferences, the results provide support for the idea, previously defended in the introduction, that video should prevail over audio to teach listening/audiovisual comprehension, or that at least both types of materials should have a similar weight.
Most informants (26 / 72.3%) responded that the visual input was “very” or “very much” useful to understand the speakers. One participant commented on item 4 of the second questionnaire the following: “Pienso que la imagen es una ayuda para la comprensión” (“I think the image aids comprehension”, author’s translation). The results are in line with some of the findings in Sueyoshi and Hardison (2005, p. 682) and Wagner (2010b, p. 287). Sueyoshi and Hardison asked 42 English as a second language (ESL) students if they agreed with three different statements: “it is easier to understand English when I can see the speaker’s face”, “it is easier to understand English when I can see the speaker’s gestures” and “it is easier to understand English conversations on TV than on the radio” (p. 697). Participants’ answers revealed agreement with the three utterances. 14 of the 42 ESL learners were exposed to a video where it was possible to see the speakers’ gestures and face. Afterwards, they were asked if they agreed that watching the speaker’s gestures, on the one hand, and watching the speaker’s face, on the other hand, helped their understanding. Globally, students sympathized with both ideas. Wagner (2010b) administered a post-test questionnaire to 56 ESL test-takers to explore their attitudes towards the use of video. The survey consisted of seven 5-point Likert-type items. Globally, participants’ attitudes were positive. One of the items requested informants to express agreement or disagreement towards the statement “being able to see the video made the test easier” (p. 286) and most informants “agreed”. Other authors found evidence that learners had a negative view of the video in listening/audiovisual comprehension tests (Alderson et al., 1995; Coniam, 2001; Suvorov, 2008). As mentioned in the introduction, those mixed results could be due to several reasons (the conflict in visual attention between watching the video and completing the task, the focus of the questions, the complexity of the task and the text, etc.).

<table>
<thead>
<tr>
<th>Table 8. How do students practise listening/audiovisual comprehension outside the classroom?</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Percent</strong></td>
</tr>
<tr>
<td>Listening to audio recordings intended for language learning</td>
</tr>
<tr>
<td>Listening to audio recordings intended for native speakers of</td>
</tr>
</tbody>
</table>
The most frequent ways to practise listening/audiovisual comprehension outside the classroom were listening to audio recordings and watching video materials intended for native speakers. The results can be related to a study on audiovisual materials by González-Vera and Hornero Corisco (2016). The authors asked 37 English as a foreign language learners what materials they used to practise listening outside the language classroom. Among the five options provided (Podcasts, video clips, DVDs, CDs and other), the most chosen categories were video clips (30%), CDs (28%) and DVDs (23%). In addition, the results revealed that students preferred authentic materials rather than materials intended for language learning. This is in line with the value of realia in some current language teaching methods, such as communicative language teaching and task-based language teaching. As for the results regarding “talking to native speakers in Spanish” and “talking to non-native speakers in Spanish”, they seem logical because, on the one hand, Spanish as a foreign language was being learnt and, on the other hand, in Coimbra there were opportunities to interact with Spanish exchange students, Spanish tourists and Spanish residents.

3.2. Research question 2

The second research question was Does the use of questions embedded within the video in the form of subtitles and synchronized with the relevant fragments in a FL audiovisual comprehension test facilitate test-takers’ performance? Do the test-takers of the experimental group score higher or lower than the test-takers of the control group?

As described in Casañ Núñez (2016b), the scores of both versions of the test were subject to descriptive statistics, analyses of facility, discrimination and reliability, and correlations with two tasks from a listening test. Overall, the results showed that the instruments were adequate for the research purposes for which they were designed. Just to mention some aspects, Cronbach's alpha was .650 in the traditional test and .705 in the experimental one (according to Suhr & Shay [2009], alphas over .60 are acceptable for instruments developed for research purposes); in both tests, the mean inter-item correlation for the items was in the .2 to .4 range (as recommend by Briggs & Cheek [1986]); and the audiovisual comprehension tests were significantly correlated in a positive way with two tasks from a listening exam (experimental test: $r_s = .592; p < .01$; traditional test: $r_s = .833; p < .01$).

<table>
<thead>
<tr>
<th>Test</th>
<th>N</th>
<th>Min</th>
<th>Max</th>
<th>M</th>
<th>SD</th>
<th>Mdn</th>
<th>Skewness</th>
<th>SE</th>
<th>Kurtosis</th>
<th>SE</th>
</tr>
</thead>
<tbody>
<tr>
<td>Trad.</td>
<td>22</td>
<td>1.00</td>
<td>7.00</td>
<td>5.182</td>
<td>1.652</td>
<td>5.500</td>
<td>-1.016</td>
<td>.491</td>
<td>.731</td>
<td>.953</td>
</tr>
<tr>
<td>Exp.</td>
<td>19</td>
<td>2.00</td>
<td>7.00</td>
<td>5.053</td>
<td>1.779</td>
<td>5.000</td>
<td>-1.618</td>
<td>.524</td>
<td>-.814</td>
<td>1.014</td>
</tr>
</tbody>
</table>

Abbreviations: Trad. = traditional. Exp. = Experimental.
Table 10. Test Statistics*

<table>
<thead>
<tr>
<th></th>
<th>Mann-Whitney U</th>
<th>Wilcoxon W</th>
<th>Z</th>
<th>Asymp. Sig. (2-tailed)</th>
<th>Exact Sig. (2-tailed)</th>
<th>Exact Sig. (1-tailed)</th>
<th>Asymp. Sig. (1-tailed)</th>
<th>Point Probability</th>
</tr>
</thead>
<tbody>
<tr>
<td>Test</td>
<td>203.500</td>
<td>393.500</td>
<td>-.147</td>
<td>.883</td>
<td>.885</td>
<td>.444</td>
<td>.009</td>
<td></td>
</tr>
</tbody>
</table>

*Grouping Variable: type of test (traditional test, experimental test).

As can been seen in Table 9, the mean scores were very similar in both tests, suggesting that questions embedded within the video did not have an effect on test performance. Since the distributions of the test scores were not normally distributed either on the experimental test ($S-W = .874$, $df = 19$, $p = .017$) or the traditional test ($S-W = .888$, $df = 22$, $p = .017$), it was not appropriate to use an independent-samples t-test. Instead, the Mann–Whitney test was employed (see Table 10). The results revealed no statistically significant difference between the test scores obtained by the control group ($Mdn = 5.500$) and the experimental group ($Mdn = 5.000$), $U = 203.500$, $z = -.147$, $p$ (exact, two-tailed) = .885, $r = -0.023$, which indicated that the subtitled questions did not have any impact on test performance. The results did not confirm the hypothesis that the experimental group would outperform the control group thanks to the help provided by the questions embedded in the video. One can think of two different explanations for that. First, perhaps the hypothesis was wrong. It could be the case that questions embedded in the video as subtitles do not constitute either a help or a hindrance when they are combined with questions on paper because learners do not pay attention to them. Second, the hypothesis and the reasoning behind it were right but they did not hold in the study for some reason. It could be that the groups were not equivalent. Although they had a similar profile, participants did not complete a pre-test to determine their level of audiovisual comprehension in Spanish.

3.3. Research question 3

The third research question investigated test-takers’ attitudes towards the use of questions embedded in the video as subtitles. Research question 3.1. inquired to what extent test-takers in the experimental group agreed or disagreed that embedded questions helped them to complete the tasks. A helpfulness scale (items 1, 2.1, 2.2, 2.3, 2.4, 5, 6.1, 6.2 and 6.3) was employed to try to answer this matter. To measure its reliability, Cronbach’s alpha, corrected item-total correlation (CITC) and Cronbach’s alpha if item deleted (CAID) values were calculated. Cronbach’s alpha was .778. According to Cohen et al. (2011: p. 640), alpha values between .70 and .80 indicate that the scale is “reliable”. All but one CITC values were above .30 (between .365 and .763), as recommend by De Vaus (2002) and Pallant (2011). Item 6.3 had a CITC level of .254, which is acceptable for Henning (1987, cited in Green, 2013). CAID values were between .709 and .781. In other words, none of the items would substantially increase the reliability if they were deleted. As the analysis was satisfactory, a new variable containing the sum of the component items was created. For that purpose, each “disagreement” was counted as 1 point, each “neither agree nor disagree” as 2 and each “agree” as 3. The scale score was checked for outliers and two were identified (see Figure 6). However, as Osborne and Overbay (2004, p. 3) point out, “there is a great deal of debate as to what to do with identified outliers”. Following Larson-Hall’s (2010) proposal, statistics with and without them were calculated (see Tables 11 and 12).
The minimum score for the scale could be 9 (for answering "disagree" to all 9 items) and the maximum 27 (for responding "agree" to all 9 items). Therefore, the higher the score, the greater the agreement that embedded questions helped learners complete the test. The mean of the scale with outliers (\( M = 20.42 \)) suggests that, overall, participants considered that imprinted questions were useful. The mean of the scale without outliers (\( M = 21.47 \)) revealed the same fact even more clearly. The results support the hypothesis and they are in line with those obtained in a previous study with SFL learners (Casañ Núñez, 2016a). In that investigation the mean of the scale was 11.50 out of 15 (SD = 2.565).

Research question 3.2 explored whether any subtitled question was consistently considered more or less helpful than the others by the experimental group. The results of the Friedman test (see Table 13) indicated that there were statistically significant differences in the ranking of the embedded questions, \( \chi^2 (6, n = 19) = 16.653, p = .011 \). Two embedded questions had particularly low mean ranks: 4 and 7. They differed

Figure 6. Boxplot of the scale scores.

<table>
<thead>
<tr>
<th>Table 11. Descriptive Statistics for the scale with two outliers</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
</tr>
<tr>
<td>Scale</td>
</tr>
<tr>
<td>Valid N</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Table 12. Descriptive Statistics for the scale without outliers</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
</tr>
<tr>
<td>Scale</td>
</tr>
<tr>
<td>Valid N</td>
</tr>
</tbody>
</table>
from the others in two ways. First, they measured the ability to recognize feelings (see table 3) and second, they did not have a focalising character: they implied global comprehension and the subtitles appeared at the end of the videos. Wilcoxon Signed Rank tests were conducted to follow up the findings. A Bonferroni adjusted alpha value was used to reduce the chances of obtaining false-positive results. As recommended by Field (2009) and Pallant (2011), selective comparisons were chosen to keep the alpha at a manageable level. Two Wilcoxon Signed Rank tests were carried out. Thus, the alpha level of significance was .025. Embedded question 4 (focusing on recognizing feelings, not focalised in the pertinent scene, level A2/B1) was compared with embedded questions 2 (focusing on identifying general ideas, synchronized with the relevant fragment, level A2/B1) and 3 (focusing on extracting specific information, synchronized with the pertinent fragment, level A2). Wilcoxon tests revealed that embedded question 4 was ranked significantly lower than question 2, \( z = -2.311, p \) (exact, two tailed) = .024, \( r = -.375 \), and question 3, \( z = -2.443, p \) (exact, two tailed) = .018, \( r = -.396 \). These facts could be due to the lack of focalization of embedded question 4. However, it would be necessary to ask test-takers to disclose how they evaluated the helpfulness of the imprinted questions.

<table>
<thead>
<tr>
<th>Table 13. Results from the Friedman test.</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Ranks</strong></td>
</tr>
<tr>
<td><strong>Questionnaire item / Subtitled question</strong></td>
</tr>
<tr>
<td>2.1. / 1</td>
</tr>
<tr>
<td>2.2. / 2</td>
</tr>
<tr>
<td>2.3. / 3</td>
</tr>
<tr>
<td>2.4. / 4</td>
</tr>
<tr>
<td>6.1. / 5</td>
</tr>
<tr>
<td>6.2. / 6</td>
</tr>
<tr>
<td>6.3. / 7</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Test Statistics*</th>
</tr>
</thead>
<tbody>
<tr>
<td>N</td>
</tr>
<tr>
<td>Chi-Square</td>
</tr>
<tr>
<td>df</td>
</tr>
<tr>
<td>Asymp. Sig.</td>
</tr>
</tbody>
</table>

*Friedman Test

Research question 3.3. investigated whether students in the experimental group were in favour of having the questions embedded in the video in the audiovisual comprehension test of the final exam. In order to address that matter, test-takers were asked directly after completing each task of the test (items 3 and 7 of the second questionnaire). The answers were virtually the same both times: most informants were in favour and very
few against this (see Tables 14 and 15). These results suggest that students considered that embedded questions were useful.

<table>
<thead>
<tr>
<th>Table 14. Item 3 (task 1). Are you in favour of having the questions embedded in the image (in addition to having them on paper) in the audiovisual comprehension test of the final exam?</th>
</tr>
</thead>
<tbody>
<tr>
<td>Frequency</td>
</tr>
<tr>
<td>No</td>
</tr>
<tr>
<td>I am not sure</td>
</tr>
<tr>
<td>Yes</td>
</tr>
<tr>
<td>Total</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Table 15. Item 7 (task 2). Are you in favour of having the questions embedded in the image (in addition to having them on paper) in the audiovisual comprehension test of the final exam?</th>
</tr>
</thead>
<tbody>
<tr>
<td>Frequency</td>
</tr>
<tr>
<td>No</td>
</tr>
<tr>
<td>I am not sure</td>
</tr>
<tr>
<td>Yes</td>
</tr>
<tr>
<td>Total</td>
</tr>
</tbody>
</table>

4. Conclusion

Listening, watching, reading and writing at the same time in a foreign language is a cognitively demanding task. This paper is part of wider research which investigates the use of audiovisual comprehension questions imprinted in video images in the form of subtitles and synchronized with the relevant fragments, for the purpose of language learning and testing. Compared to viewings where the task is available only on paper, this technique may provide some benefits. Among them, it could reduce the conflict in visual attention between watching the video and completing the task, by spatially and temporally approximating the questions and the relevant scenes. The procedure is mainly intended for students with a low proficiency level.

This pilot multimethod study (Morse, 2003) investigated for the first time whether questions embedded in videos as subtitles had any impact on FL learners’ audiovisual comprehension test performance. The results suggest that they do not have any effect on test performance. Test-takers’ attitudes towards this technique are positive, however. Overall, participants in the experimental group agree that the imprinted questions help them to complete the tasks. Furthermore, most of them are in favour of having the questions embedded in the video in the audiovisual comprehension test of the final exam. Test-takers’ opinions can be paralleled to studies that research experts’, SFL students’ and SFL teachers’ views about this methodology (Casañ Núñez, 2015a, 2016a, in press-b). Globally, all three groups have positive opinions. Experts agree or strongly agree that the technique can be useful for the teaching of listening/audiovisual comprehension, and that it can provide various benefits compared to viewings where the activity is available only on paper; among them, that it can minimize the conflict in visual attention between watching a video and completing a task at the same time, and
that it helps FL students to focus their attention towards the viewing objectives. SFL teachers think that embedded questions are beneficial for FL learning, and they coincide with the experts in some of the advantages. SFL students agreed that imprinted questions helped them to complete a testing task; besides, their comments imply that subtitles focalise attention in the relevant moments, and that they minimize the conflict of visual attention. All in all, although these studies are limited and further empirical research is needed, the positive opinions suggest that this technique has potential benefits for FL learning and testing.

The current study has a number of shortcomings. First, SFL students were not selected by probabilistic sampling. As Dörnyei (2007) points out, this weakness is present in most experimental studies in the social sciences. Second, treatment diffusion constituted a potential threat to internal validity because the groups took the audiovisual test consecutively in the same classroom. Two circumstances reduced the possibility of the exchange of information: (a) there was little time for it to happen, since test-takers took the tests consecutively and they had different class schedules; and (b) in principle, there was no obvious interest in knowing the test content, as it had no impact on the final grade. A further limitation consists in the impossibility of guaranteeing that the groups were of equal ability. Although they were similar in many aspects and there were no statistically significant differences between the groups’ performance on two tasks from a listening test, an audiovisual comprehension test to find out the participants’ level of audiovisual comprehension in SFL was not administered.

This leads us to some directions for future research. First, it is important to investigate whether embedded questions have an impact on FL learners’ viewing behaviour with regard to the video image. Compared to viewings where the task is available only on paper, imprinted questions may reduce the conflict in visual attention between watching the video and completing the task, by spatially and temporally approximating the questions and the relevant fragments. Based on this, it is hypothesized that embedded questions may increase the amount of time devoted to viewing the video by the learner. Second, it would be beneficial to replicate this study with a larger sample and compare the results. Third, it would be useful to carry out similar studies with tests for different target language use domains, other text types and other text lengths. Lastly, as this work explored embedded questions in a testing situation, further research is needed to investigate this technique in a learning context.

Dedication

This article is dedicated to my father.

References


Appendix

**Preguntas integradas como subtítulos: la opinión del estudiante (original version)**

El presente cuestionario tiene por finalidad conocer qué piensas sobre el uso de las preguntas de comprensión integradas en el vídeo en forma de subtítulos. No existen respuestas correctas ni incorrectas porque estás expresando tu opinión. La información que proporciones se tratará de forma confidencial y solo se utilizará con fines académicos.

**Tarea 1. Fragmento de la película Ópera prima.**

1. Sobre la base de tu experiencia global completando la tarea, ¿en qué medida estás de acuerdo o en desacuerdo con que las preguntas subtituladas te han ayudado? Señala la respuesta con un círculo.
   a) En desacuerdo  
   b) Ni de acuerdo ni en desacuerdo  
   c) De acuerdo

2. De forma individual, ¿en qué medida estás de acuerdo o en desacuerdo con que las preguntas subtituladas te han ayudado? Completa la tabla poniendo equis (X).

<table>
<thead>
<tr>
<th></th>
<th>En desacuerdo</th>
<th>Ni de acuerdo ni en desacuerdo</th>
<th>De acuerdo</th>
</tr>
</thead>
<tbody>
<tr>
<td>2.1. ¿Cómo se llama la chica?</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2.2. ¿De qué hablan?</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2.3. ¿Cuál es el teléfono de la chica?</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2.4. ¿Qué sentimientos puede tener el chico por la chica?</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

3. ¿Estás a favor de que en la prueba de comprensión audiovisual las preguntas aparezcan en la imagen (además de en papel)?
   a) No  
   b) No estoy seguro  
   c) Sí

4. A continuación tienes un espacio en el que puedes comentar cualquier aspecto de la actividad. Puedes responder en español o portugués.

**Tarea 2. Fragmento de la película Los peores años de nuestra vida**
5. Sobre la base de tu experiencia global completando la tarea, ¿en qué medida estás de acuerdo o en desacuerdo con que las preguntas subtituladas te han ayudado? Señala la respuesta con un círculo.
   a) En desacuerdo
   b) Ni de acuerdo ni en desacuerdo
   c) De acuerdo

6. De forma individual, ¿en qué medida estás de acuerdo o en desacuerdo con que las preguntas subtituladas te han ayudado? Completa la tabla poniendo equis (X).

<table>
<thead>
<tr>
<th></th>
<th>En desacuerdo</th>
<th>Ni de acuerdo ni en desacuerdo</th>
<th>De acuerdo</th>
</tr>
</thead>
<tbody>
<tr>
<td>6.1. ¿De qué temas habla el chico pelirrojo?</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>6.2. ¿Cómo se llama la chica?</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>6.3. Considerando todo el fragmento, ¿qué sentimientos puede tener el chico pelirrojo por la chica? Justifica brevemente la respuesta.</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

7. ¿Estás a favor de que en la prueba de comprensión audiovisual las preguntas aparezcan en la imagen (además de en papel)?
   a) No
   b) No estoy seguro
   c) Sí

8. A continuación tienes un espacio en el que puedes comentar cualquier aspecto de la actividad. Puedes responder en español o portugués.

Questions embedded as subtitles: students’ opinion (English translation)

This questionnaire aims to know what you think about the use of audiovisual comprehension questions embedded in the video in the form of subtitles. There are no right or wrong answers because you are expressing your opinion. The information you provide will be treated confidentially and will only be used for academic purposes.

Task 1. Fragment from the film Ópera prima.

1. Based on your overall experience completing the task, to what extent do you agree or disagree that the subtitled questions have helped you? Mark your answer with a circle.
   a) Disagree
   b) Neither agree nor disagree
   c) Agree

2. Individually, to what extent do you agree or disagree that subtitled questions have helped you? Complete the table with “X”.

<table>
<thead>
<tr>
<th></th>
<th>Disagree</th>
<th>Neither agree nor disagree</th>
<th>Agree</th>
</tr>
</thead>
<tbody>
<tr>
<td>2.1. What is the girl’s name?</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
2.2. What are they talking about?  

2.3. What is the girl’s telephone number?  

2.4. What feelings may the boy have for the girl?  

3. Are you in favour of having the questions embedded in the image (in addition to having them on paper) in the audiovisual comprehension test of the final exam?  
   a) No  
   b) I am not sure  
   c) Yes  

4. Next there is a space where you can comment on any aspect of the activity. You may reply in Spanish or Portuguese.  

**Task 2. Fragment of the film Los peores años de nuestra vida.**  

5. Based on your overall experience completing the task, to what extent do you agree or disagree that the subtitled questions have helped you? Mark your answer with a circle.  
   a) Disagree  
   b) Neither agree nor disagree  
   c) Agree  

6. Individually, to what extent do you agree or disagree that subtitled questions have helped you? Complete the table with “X”.  

<table>
<thead>
<tr>
<th></th>
<th>Disagree</th>
<th>Neither agree nor disagree</th>
<th>Agree</th>
</tr>
</thead>
<tbody>
<tr>
<td>6.1. What is the redheaded boy talking about?</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>6.2. What is the girl’s name?</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>6.3. What feelings may the red-haired boy have for the girl?</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

7. Are you in favour of having the questions embedded in the image (in addition to having them on paper) in the audiovisual comprehension test of the final exam?  
   a) No  
   b) I am not sure  
   c) Yes  

8. Next there is a space where you can comment on any aspect of the activity. You may reply in Spanish or Portuguese.
Research paper

Profiling language learners in hybrid learning contexts: Learners’ perceptions

Pekka Lintunen*, Maarit Mutta** and Sanna Pelttari***
University of Turku, Finland

*pekka.lintunen@utu.fi | **maarit.mutta@utu.fi | ***sanna.pelttari@utu.fi

Abstract

This article discusses formal and informal foreign language learning before university level. The focus is on beginning university students’ perceptions of their earlier learning experiences, especially in digital contexts. Language learners’ digital competence is a part of their everyday lives, but its relationship to learning in and outside educational settings is still relatively seldom studied. The article discusses learning in formal and informal (i.e., hybrid) contexts and digital learning profiles – that is, a learner’s own personalized style in acquiring language competence by creating affordances in personalized digital or mobile learning environments – in primary and secondary education identified in a language learning survey. The results are based on an online survey sent to all beginning university students majoring in languages at a Finnish university (N= 87/192), which was complemented by a short narrative task (N=47) a few months later focusing on earlier education and the use of language learning technologies. The results suggest that the use of technologies seems to differ between extramural and in-school language learning. The learners were well aware of various possibilities to create affordances for learning, and their own involvement increased with age. Most participants had positive attitudes towards the use of technologies to enhance language learning, but critical views emphasized the importance of inspiring contact teaching. Three different digital learning profiles were identified: a digiage learner, a hybrid learner, and an in-school learner. These can be useful when planning differentiated foreign language instruction.

Keywords: Computer-assisted language learning, social media, learner behaviour, attitude.

1. Introduction

Foreign language learners acquire the target language from many sources. Languages are understood as dynamic means of communication, and the learning process is strengthened by naturalistic exposure and communicative situations requiring creative language use with emerging skills. Language learning can be implicit (i.e., without attention) or explicit (i.e., conscious), and often these two processes are intertwined and inseparable. Language acquisition takes place in hybrid environments, i.e., formal and informal contexts. Moreover, digital technologies and learning environments are part of the everyday lives of the language learners of today (Erstad, 2010; Jalkanen & Talas, 2015; Mutta, Lintunen, Ivaska & Peltonen, 2014; Palmgren-Neuvonen, Jaakkola & Korkeamäki, 2015; Pirainen-Marsh & Tainio, 2009). This has affected the contexts of foreign language learning fairly recently (Blin & Jalkanen, 2014; Jalkanen, 2015; Thorne & Reinhardt, 2008).
Extramural language learning studies can be quantitative or qualitative, for instance, with learners using Facebook (Mitchell, 2012), telecollaboration via email (Schenker, 2012) or voice blogs (Sun, 2012) to facilitate their language and cultural learning. Researchers have also been interested in learners’ out-of-school activities, such as gaming, and their impact on language learning (Sundqvist & Sylvén, 2014; see also Sundqvist, 2016; Thomas & Peterson, 2014). These extramural activities seem to increase learners’ motivation and, thereby, indirectly promote language learning. Extramural language learning can also take place via traditional ways, such as trips abroad and interaction with foreigners.

The younger generations are commonly considered as digital natives, who use various technologies frequently and fluently and are used to certain everyday interactions in the social media. However, they may not be so familiar with technologies supporting digital language learning, the uses of social media for learning purposes, and teaching environments utilizing different technologies (Ilomäki, Taalas & Lakkala, 2012; Laakkonen, 2015; Mutta, Pelttari, Lintunen & Johansson, 2014; Williams, Abraham & Bostelmann, 2014). There might also be vast discrepancies between in and out-of-school use of technology (Jalkanen, 2015; Keltikangas-Järvinen, 2015). We also have to acknowledge that not all young learners can be categorized as digital natives: although born during the digital age, their skills might not match the implications of the concept. Instead, the concept can be used to characterize certain individual learners while admitting that some might not feel confident or comfortable using various technologies for formal or social purposes (Benini & Murray, 2014; Mutta, Pelttari, Lintunen, et al., 2014; Thomas, 2011; Thomas & Peterson, 2014). The variation may be related to, for example, personality, learning experiences, and different digital learning profiles.

Learning profiles and attitudes towards learning can be studied with quantitative, qualitative or mixed methods, such as surveys, interviews, and learning diaries, often depending on the size of the study (e.g., Eloranta & Jalkanen, 2015; Ilomäki et al., 2012; Jalkanen & Taalas, 2013; Sundqvist & Sylvén, 2014; Williams et al., 2014). This study presents findings from a survey aimed at beginning university students majoring in languages at a university in Southern Finland complemented by a short narrative task. It discusses learner perceptions of and attitudes towards extramural language learning, learners’ knowledge and use of technologies, and their digital learning profiles before tertiary level language education.

2. Literature review

When it comes to defining new technologies and language learning, terminology is varied (e.g. computer-enhanced language learning, Web 2.0 and L2 learning, technology-enhanced language learning, mobile-assisted language learning), and discussions about the most appropriate term are ongoing (Healey, 2016). Jalkanen (2015) talks about technology-rich environments, whereas Wang and Vásquez (2012: 413), for example, use the concept Web 2.0 technologies, but Thomas and Peterson (2014: i) suggest that “social media” is sometimes used as its substitute term. Nevertheless, concepts related to new technologies have become intuitive and fuzzy needing further discussion. For practical reasons, we gave examples of different technologies in our survey without categorizing them into old or new.

As a ubiquitous use of various technologies characterizes our world today, researchers have attempted to define the new generations in this information and communication technology (ICT) filled world. Prensky (2001: 1) called them digital natives (as opposed to digital immigrants), who were “all ‘native speakers’ of the digital language of computers” and used various technologies frequently with great ease in everyday interaction (see also Williams et al., 2014: 30). This view has been challenged by several scholars based on classroom studies (e.g., Benini & Murray, 2014; Williams et al., 2014; see also Thomas, 2011).
Williams et al. (2014) synthesized six survey-driven studies from different countries to motivate their two large-scale surveys on undergraduate foreign language learners’ perceptions and use of digital tools at a large university in the USA. The researchers concluded that the learner’s age is not the primary determiner for digital practices: even if the learners are familiar with various technologies, they do not unanimously identify themselves as digital natives. More importantly, they may not transfer their knowledge and skills of social media to support language learning.

Benini and Murray (2014) conducted a study on attitudes and the uses of digital technologies by monitoring and interviewing students and teachers in two secondary education institutions in Ireland. They suggested that students and teachers are very similar in their everyday use of technologies, but the types of activities they are involved in seem to be different, especially in an educational context; students considered technologies stimulating but not essential for foreign language learning. As mentioned above, several recent studies have discussed the dissimilarities between in and out-of-school use of technologies (e.g., Ilomäki et al., 2012; Laakkonen, 2015; Mutta, Pelttari, Lintunen, et al., 2014; Williams et al., 2014). Studies have shown, for instance, that students may be fairly used to sending emails, navigating online, and using social networking for personal purposes, whereas the same tools are less used in classroom environments depending on the resources and facilities in their school (Benini & Murray, 2014; Williams et al., 2014; see also Conole, 2008).

According to the Programme for International Student Assessment (PISA) (1), Finland has been among the top countries in learning outcomes for several years. Admittedly, however, this assessment has focused on subjects such as mathematics and science or first language reading skills, and not assessed the use of technology in schools. Finland has the reputation of being a technologically advanced country with the success of Nokia and, more recently, thanks to the digital game industry such as the Angry Birds franchise. The Finnish National Board of Education deduced in 2013 based on a European survey (Survey of Schools: ICT in Education) (2) that although Finnish schools are equipped with the latest technology, the active use of ICT in teaching is far behind many other countries.

In Finland, in and out-of-school language learning environments seem to create two separate spheres in the lives of young learners: the digital life outside school is rich, whereas in the school environment digital skills may have been long neglected. Based on their vast survey (N=1700 ninth grade students and 740 teachers), Ilomäki et al. (2012) discovered a clear gap between in and out-of-school activities using technologies. Furthermore, the teachers’ free time technology use was more instrumental (e.g., seeking for information), whereas students used technologies to spend time (e.g., by playing online games; see also Jalkanen, 2015). The schools, educators, and teachers, however, pay more attention to this discrepancy to bring these two realities closer; there are several recent studies on the use of technologies in different learning contexts in Finland (e.g., Ilomäki et al., 2012; Jalkanen & Taalas, 2015; Jokinen & Vaarala, 2015; Vaarala, Johansson & Mutta, 2014; Vaarala & Lehtonen, 2015).

When the effects of extramural foreign language learning contexts have been studied in Finland, boys have been found to outperform girls in English language tests because they have more access to informal learning, for instance, by playing video games online (Ilomäki et al., 2012; Uusikoski, 2011; see also Piirainen-Marsh & Tainio, 2009). Sundqvist and Silvén’s (2014) study revealed similar results in Sweden: among young learners (aged 10-11; N=76) the frequent gamers, who had more extramural training, were all male, and their English language proficiency and confidence was evaluated higher in communicative situations (see, however, Sundqvist, 2016). Extramural activities create more affordances to promote (mostly implicit) language learning, but they do not primarily focus on language learning.
Conole (2008) encourages listening to learner voices in her in-depth case study on language learners’ use of technologies. We study the learners’ voices and describe them as digital learning profiles, which refer to learners’ own personalized styles in acquiring language competence by creating affordances in personalized digital or mobile learning environments. This digital learning profile comprises in and out-of-school use of technology, but the latter requires more effort and personal investment by the learner. For their part, learning paths are often studied as part of foreign language classroom learning to understand learners’ linguistic backgrounds, aims, motivations, and needs towards formal education and its functional purpose for future interaction in various social contexts (Eloranta & Jalkanen, 2015). Different paths can also reveal the strategies used to complete a specific linguistic action (Mutta, Pelttari, Salmi, Chevalier & Johansson, 2014). E-learning or a digital learning path is a relatively new concept, which does not yet have a well-established definition. It can be related to Personal Learning Environment (PLE), which refers both to creating content on the Internet and to personal experiences promoting learning in different contexts (Attwell, 2007; Guth, 2009; Laakkonen, 2015). Jalkanen and Taalas (2015: 183) give an overview of multimodal foreign language pedagogies, which are mostly Computer-Assisted Language Learning (CALL) studies, and recognize the need to study how language learners’ digital paths develop from one educational level to the next. However, studying paths requires a longitudinal approach with multifold observations. Instead, we focus on profiles based on a retrospective survey to study learners’ perceptions on their use of technologies in hybrid contexts. In her study on metacognitive knowledge development and language learning in web-based distance learning contexts, Fincham (2015) observed learner profiles based on an initial survey and an interview. Her survey consisted of questions, for instance, on learners’ previous experience with language learning and various multimedia tools in different contexts.

The objective of the present study was to gain new insights into extramural learning and the use of digital technologies before tertiary language education in Finland and learners’ perceptions of their use in foreign language learning and teaching contexts. Although other surveys have been conducted in Finland and elsewhere, there are still not many studies on hybrid or extramural foreign language learning and learner opinions and attitudes. Furthermore, tracing learners’ digital profiles is a recent area of interest. The study contributes to the knowledge of the distribution of utilizing digital and non-digital resources and the reality of so-called digital natives at the beginning of tertiary education. The understanding of learners’ digital profiles enables educators to recognize learners’ personal needs and learning styles to promote the creation of personalized digital learning environments.

The study focuses on the following research questions:
1. What kind of extramural language learning is recognized and in which context?
2. What kinds of digital technologies have been used?
3. What kinds of attitudes are revealed towards the use of various technologies in foreign language learning?
4. What kinds of digital learning profiles can be identified among the participants?

3. The present study
An online Webropol survey was administered at the beginning of the 2015 autumn semester to all first-year students (N=192) majoring in languages at a university in Southern Finland. The response rate was 45.8, and altogether 87 students provided answers for the survey. The survey, adapted from the technology implementation questionnaires prepared by the Centre for the Study of Learning and Performance (CSLP) (3), had 36 (mostly multiple-choice) questions in three sections: 1. Demographic information (13 questions, e.g., background), 2. Questions on extramural language learning and use of technologies (17 questions separately for different levels of education), and 3. Attitude questions (6 questions focusing on the advantages and disadvantages of technologies). To strengthen the analysis, we asked students to write
a short narrative (N=47) some months later to allow students to elaborate more freely on their earlier experiences as language learners and the use of language learning technologies. These answers were used to support the identification of different learning profiles. All numerical information is based on the larger survey. After a brief account of background information, we will discuss the general use of technologies, learners’ extramural foreign language activities before tertiary education, and finally, analyze learner attitudes. This small-scale study is mainly of a qualitative nature. The quantitative analyses were made with the Webropol statistical tool.

The participants were mostly female (female 84.9%, male 15.1%), which was expected as language students at university level are predominantly female in Finland. The youngest participants were 19 years old, and seven were born in 1980 or earlier. Most participants (56%) were 20 years old or younger. All participants had passed an entrance examination with strict selection criteria to enter university, which means that their target language proficiency was quite high (especially in English which is the most common first foreign language in Finland) and they can be considered highly motivated to learn foreign languages. Accordingly, their responses should give a fairly accurate account of the spectrum of technologies used for extramural language learning. Over 93% of the participants had Finnish as their first language; the other first languages mentioned were Swedish, Russian, English, and Italian.

Among the participants, the most common major subject at university was English: 37 participants studied English as their major (43%), followed by Finnish (17.4%) and French (12.8%); the other mentioned languages were Swedish (8.1%), German (5.8%), Spanish (4.7%), Italian (4.7%), and finally, Greek, Latin and Russian (3.5%). In addition, 80.5% of the participants indicated that English was their strongest foreign language, for six it was Swedish, for five Finnish, for one Spanish or French, whereas four participants mentioned two languages (English and French, German, Spanish, or Swedish). This information illustrates well the fact that learners are more frequently exposed to English in Finland than to other languages, even including Swedish, which is the other official language.

4. Survey results

The questions on extramural language learning and use of technologies included 6 close-ended and 11 open-ended questions separately for different educational levels. We also asked about the use of technologies in general at the beginning of tertiary education. Most participants (95.4%) reported using the Internet daily, email (89.7%), instant messaging (WhatsApp, 82.8%), Facebook (72.4%) and YouTube (51.7%). In addition, watching video clips online (44.8%) and Instagram (41.4%) were common. The technologies that were relatively unknown (i.e., never used) were MySpace (94.3% of the participants), virtual worlds (e.g., Second Life, 93.1%), creating video blogs (vlogs) (93.1%), social bookmarking web service (e.g., Delicious, 93.1%), LinkedIn (90.8%), Flickr (87.4%), and editing Wikipedia (87.4%). These results show that the participants mostly used technologies when using the Internet and communicating with other users, whereas many other resources were relatively unknown to them (cf., Benini & Murray, 2014; Ilomäki et al., 2012; Laakkonen, 2015; Mutta, Pelttari, Lintunen, et al., 2014; Williams et al., 2014).

In the following sections, we will discuss extramural language use that supports language learning before tertiary education, learners’ experiences of the technologies used in different contexts, their attitudes towards using technologies in language learning and teaching, and finally, we will discuss digital learning profiles.

4.1. Extramural language use before tertiary education

The questionnaire focused on three educational levels: primary school, secondary school, and upper secondary school. We asked in which extramural contexts the learners used foreign languages (23 contexts were given and one open-ended question), which target language was used, and how much the students thought they
had learnt in extramural contexts. Table 1 illustrates the most frequent (> 40%) activities mentioned including different languages at different stages.

**Table 1.** Extramural language activities in primary, secondary and upper secondary school.

<table>
<thead>
<tr>
<th>Rank</th>
<th>Activity</th>
<th>%</th>
<th>Rank</th>
<th>Activity</th>
<th>%</th>
<th>Rank</th>
<th>Activity</th>
<th>%</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Watching foreign language TV series/films (subtitled)</td>
<td>87.4</td>
<td>1</td>
<td>Listening to music</td>
<td>92.0</td>
<td>1</td>
<td>Watching foreign language TV series/films (subtitled)</td>
<td>89.7</td>
</tr>
<tr>
<td>2</td>
<td>Listening to music</td>
<td>87.4</td>
<td>2</td>
<td>Watching foreign language TV series/films (not subtitled)</td>
<td>88.5</td>
<td>2</td>
<td>Listening to music</td>
<td>89.7</td>
</tr>
<tr>
<td>3</td>
<td>Reading song lyrics</td>
<td>64.4</td>
<td>3</td>
<td>Reading song lyrics</td>
<td>80.5</td>
<td>3</td>
<td>Watching foreign language TV series/films (not subtitled)</td>
<td>88.5</td>
</tr>
<tr>
<td>4</td>
<td>Trips abroad</td>
<td>48.3</td>
<td>4</td>
<td>Trips abroad</td>
<td>72.4</td>
<td>4</td>
<td>Reading song lyrics</td>
<td>82.8</td>
</tr>
<tr>
<td>5</td>
<td>(Online) games</td>
<td>48.3</td>
<td>5</td>
<td>Watching foreign language TV series/films (not subtitled)</td>
<td>71.3</td>
<td>5</td>
<td>Reading novels/fanfiction</td>
<td>74.7</td>
</tr>
<tr>
<td>6</td>
<td>Reading novels/fanfiction</td>
<td>63.2</td>
<td>6</td>
<td>Trips abroad</td>
<td>71.3</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>7</td>
<td>Guiding tourists</td>
<td>51.7</td>
<td>7</td>
<td>Guiding tourists</td>
<td>65.5</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>8</td>
<td>(Online) games</td>
<td>51.7</td>
<td>8</td>
<td>Contacts with foreign language friends</td>
<td>64.4</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>9</td>
<td>Contacts with foreign language friends</td>
<td>42.5</td>
<td>9</td>
<td>Summer job</td>
<td>49.4</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>10</td>
<td>Using foreign language with Finnish friends</td>
<td></td>
<td>10</td>
<td></td>
<td>47.1</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>11</td>
<td>Writing their own texts</td>
<td></td>
<td>11</td>
<td></td>
<td>46.0</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>12</td>
<td>(Online) games</td>
<td></td>
<td>12</td>
<td></td>
<td>44.8</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Table 1 shows that the extramural foreign language activities became more varied when moving from primary to secondary school, the most frequent activities among learners being watching subtitled foreign language TV series and/or films and listening to foreign language music. This tendency continued in upper secondary school. At this level, watching foreign language TV series and/or films without subtitles ranked third (and fifth in secondary school), which illustrates increasing active involvement in creating affordances in extramural foreign language learning. Furthermore, in upper secondary school, it became more common to use foreign languages even with Finnish friends and when learners were writing their own texts, which implies more active learner involvement. Games were present at all educational levels, ranking 5th (48.3%), 8th (51.7%), and 12th (44.8%) in extramural activities, respectively.

In addition, in the open-ended questions, the participants mentioned watching YouTube videos, reading comics, role-playing, sending messages via messenger, online discussions, and following English vlogs as extramural foreign language practice at secondary school level. In upper secondary school the activities were even more varied: watching YouTube videos, listening to English podcasts, chatting on IRC, in (part-time) work, role-playing, listening to online radio, reading online newspapers/news, reading comics and books, going out with exchange students, European Youth Parliament activities, using social media (e.g., Instagram), and reading blogs. These results show that language learners’ extramural activities are multifold, and digital environments are only one part: learning environments are hybrid spaces (cf., Blin & Jalkanen, 2014).

In primary school, 83 learners (95.4%) used English in extramural activities; the other foreign languages mentioned were Swedish (12.6%), German (9.2%), French (8.1%), Finnish (3.5%), and Chinese, Italian, Japanese, Dutch, and Russian (1.2%). Most Finnish schools only teach English as the first foreign language in primary school, which is reflected in the answers. In secondary school, English was still the most used foreign language (93.1%), followed by Swedish (32.2%), French (15.0%), German (12.6%), Spanish (4.6%), Japanese (3.5%), and Albanian, Chinese, Finnish, and Russian (1.2%). Finally, in upper secondary school, the order stayed mostly the same, but the uses were more frequent: English (95.4%), Swedish (48.3%), French (21.8%), German (21.8%), Spanish (15.0%), Finnish (4.6%), Russian (4.6%), Italian (3.5%), Japanese (3.5%), and Chinese, Portuguese, Serbian, and Turkish (1.2%). To summarize, from primary education onwards, English was the most used language in extramural language activities, whereas the use of Swedish, the other official language in Finland, increased from 12.6% to 48.3%. These languages were followed by French, German, and Spanish, all with increasing use at the end of upper secondary school. This order corresponds to the order of the most common foreign languages taught at Finnish schools. It seems that learners used various languages in extramural activities, but English was clearly the most used language.

Furthermore, the learners’ own activity and involvement increased in extramural contexts during upper secondary school. Most learners perceived that they had learnt a lot during extramural activities, even “much more than at school”, whereas only a few learners said that they had merely “learnt something” in extramural contexts, or their motivation was “just to support language proficiency”. The learners also emphasized their own activity in different situations such as face-to-face contexts with friends or online. The participants felt that extramural language use gave them more confidence to use languages both in and out-of-school contexts.

4.2. Technological experience in hybrid contexts

When asked if they had had any experience in using language learning technologies and applications in school contexts during previous education, 47 participants (54.0%) replied that they had, whereas 40 (46.0%) said that they had not had any experience in using them. From these 47 learners, 38 answered the question “Which of these experiences have been the best and most useful” on at least one education level (see Table 2).
Table 2. Learners’ perceptions of the most useful technology-based activities in school.

<table>
<thead>
<tr>
<th>Grade</th>
<th>N of participants</th>
<th>Activities</th>
</tr>
</thead>
<tbody>
<tr>
<td>Primary school</td>
<td>7</td>
<td>word processing programs (3), online exercises, online vocabulary exercises, video games, online videos</td>
</tr>
<tr>
<td>Secondary school</td>
<td>16</td>
<td>using the interactive white board (4), online (vocabulary) games (4), watching films/videos (2), watching foreign language programs online, online dictionaries, online teaching materials and manuals, online tests and exams, searching information online, video projectors, writing texts on a computer, YouTube</td>
</tr>
<tr>
<td>Upper secondary school</td>
<td>36</td>
<td>learning platforms (6), using the interactive white board (5), online exercises (4), foreign language clips/films/videos (3), online (vocabulary) games (3), computers (2), group work/Presentation presentations (2), internet sites (2), online dictionaries (2), PowerPoint presentations by teachers (2), Quizlet (2), using tablets/IPads (2), video/data projectors (2), creating videos, distance learning class, email penpals, google drive, Kahoot, using the language laboratory, online language courses, online teaching materials and manuals, recording and listening equipment, using smart phones, word tests with Socrative, writing texts on a friendship school’s site, YouTube</td>
</tr>
</tbody>
</table>

The variety of technologically based activities increased from primary to upper secondary school, which might also be a sign of technological advancement in schools. Differences due to schools and teachers are also evident, and some answers were more related to classroom equipment than to activities performed.

The learners were also asked about the help and scaffolding they received to use technologies in general. In primary school, 101 occurrences of support were reported (39 in-school and 62 out-of-school), in secondary school 111 (47 in-school and 64 out-of-school), and in upper secondary school 118 (50 in-school and 68 out-of-school). At primary level, ICT-lessons and support by teachers were most often mentioned in the school context, whereas in extramural contexts other family members were most frequently mentioned, followed by friends and the learner’s own input. In secondary school, teachers’ support was most frequently indicated, followed by school in general and ICT lessons; in extramural contexts friends were often mentioned, then learning by oneself, and finally, support from family. In upper secondary school, support from teachers was still the most frequent in-school support, whereas ICT lessons were seldom mentioned; furthermore, learners indicated their own active involvement in school projects. In extramural contexts, the learner’s own activity was clearly the most frequently mentioned source of support followed by friends’ help; other family members’ role decreased at this level. The analysis showed that, although schools and teachers provided help with the technologies that scaffold language learning, support was mostly received in out-of-school contexts from the learners’ peers. That is, both in-school and out-of-school contexts play a role in supporting technology-rich language learning. The importance of the learner’s own input was highlighted especially at higher levels of education.

4.3. Attitudes towards using technologies in language learning and teaching.

To study learners’ attitudes towards using technologies and the learners’ readiness and open-mindedness to use them themselves, we asked our participants to respond to 18 statements (on a four-point Likert scale; totally agree, partially agree, partially disagree, and totally disagree), one yes/no-question, and answer four open-ended questions about the usefulness of technologies for language learning. Most answers were either partially agree or partially disagree.

Using technologies in teaching and learning seemed to divide the participants into two groups: for or against their usefulness. One reason for this might be that 46% of the
surveyed participants did not report any earlier experience in using technologies to support language learning, although the learners were quite familiar with technologies in their everyday lives. The results based on all participants’ answers show that the learners responded most unanimously to the statement that technologies are useful in language learning and teaching if teachers have received appropriate training in their use (89% agreed). Most learners had positive attitudes towards using technologies in learning and teaching. For instance, 83% of the participants thought that technologies had a facilitating effect on learners’ communication skills and allow future teachers to transform from information dealers to facilitators of learning. Moreover, technologies seem to diversify teaching materials and be useful even if they were used in out-of-school contexts (82%). Additionally, 78% of the participants thought that technologies support learners’ skills at all proficiency levels, allow them to use their own personal learning style, and are not too time-consuming due to technical problems. Furthermore, the participants thought that their use does not increase learners’ stress and anxiety levels (72%), but on the contrary, seem to promote interactive cooperation (69%). Over half of the participants also believed that using technologies motivates learners to engage more in the learning process, but the same number of participants was also concerned that using various technologies might harm face-to-face contact teaching (56%).

Finally, when focusing more on those participants who had previous experience with language learning technologies and applications (N=47; 54%), there were contradictory opinions. Most participants chose either the partially agree or partially disagree alternatives. Their answers mainly corroborate those of the whole group, but the more experienced learners were more of the opinion that learning how to use technologies is time-consuming (55%) and that technologies do not motivate learners to engage more in the learning process (53%). In addition, the more experienced group was quite divided about whether learners’ knowing more about computers than their teachers has negative consequences (51% vs. 49%). On the other hand, those learners who did not have previous experience with technologies in language learning thought that learning how to use them is not time-consuming (57%), their use motivates learners to engage more in the learning process (60%), and learners’ knowing more about computers than their teachers does not have negative consequences (63%).

In conclusion, the majority of the participants reacted positively towards the use of technologies to enhance language learning, but there were also some critical views to emphasize the importance of inspiring contact teaching. Social media especially was mentioned as an out-of-school activity that should not be used for language learning, although it was recognized as a potential channel for spreading information (cf., Benini & Murray, 2014; Ilomäki et al., 2012; Laakkonen, 2015; Mutta, Pelttari, Lintunen, et al., 2014; Williams et al., 2014). Technologies were viewed as complementary to good contact teaching.

4.4. Digital learning profiles

To identify digital learning profiles, we analyzed qualitatively learners’ answers on where and how they had learnt foreign languages, how much they used technologies in general, and what kinds of attitudes they had towards the use of technologies in language learning. Our approach was fairly similar to Fincham (2015), but instead of interviews we complemented our survey results with lengthier open-ended answers a few months later from 47 first year students. By using the grounded approach (e.g., Mutta, Pelttari, Salmi, et al., 2014; see also Fincham, 2015), we identified three digital learning profiles to illustrate the multifold phenomenon. The categories are tendencies and not clear-cut, but provide examples of different learner types among the generation of digital native learners. They also reflect how learners have different goals and methods for language learning. Next we will introduce three typical profiles named after learner types: a digiage learner, a hybrid learner, and an in-school learner. The
quotations below illustrating the profiles are from learners’ responses to open-ended questions.

Digiage learners have used various technologies on a daily basis especially in their out-of-school social activities. See quotations 1, 2 and 3.

1. *I have a positive attitude towards language learning technologies. If they are used well, they support learning. Especially mobile devices enable learners to carry e-books and electronic dictionary applications with them always and almost wherever. In addition, online exercises that give you immediate feedback are good means to learn languages.*

2. *My attitude is positive. We live in a world where the use of technologies is on the increase, and people have more knowhow on technologies. We are surrounded by electronic devices and the media, so using and integrating them into learning could be effortless. I think that they are a natural addition to other learning styles and equipment that support learning.*

3. *It is good that modern technologies can be used in teaching. Often people use many mobile devices in their free time so that they could also be used to support learning. You can find many dictionaries and applications online to improve your language learning.*

These digiage communicators are heavy users of social media (Facebook, WhatsApp, Instagram, Snapchat, and Twitter) and the Internet, and might use around ten different activities online daily. They have started using various technologies more outside school with age, but do not have much experience in their classroom use. For instance, the participants who wrote the quotations said that they had no experience in the use of language learning technologies at school (the participant who wrote the second quotation had tried some online exercises). The digi age learners have also often played games online. Their attitudes towards new technologies might be positive, but not entirely established. Social media have been mostly part of their out-of-school activities, but not part of formal education, perhaps due to the nature of their previous education and pedagogic choices of their former teachers.

Hybrid learners have used technologies in remarkably versatile ways in their out-of-school activities: for example, social media have been used not only for social purposes such as language learning. Quotations 4, 5 and 6 are from participants who had used many technologies in secondary education and were experienced users of social media.

4. *The use of technology usually motivates students. Even the simple tasks that would be dull and boring as grammar handouts can be surprisingly good when they are introduced with some application. However, the use of technology can also be distracting and steal the focus from the issue discussed if, for example, the application does not work like it should or if it helps learning only superficially (for instance, if a student uses an ipad application and the teacher thinks that they will learn target language words at the same time, but the application allows the student pass based on mere guesswork)"…"As a child, I learnt a lot about a foreign language via using the social media (or some earlier version of it), although it was not much used in teaching then.*

5. *Using technologies as teaching methods is not bad at all, but it all comes back to the comfort zone of the teacher. I would rather take down notes by hand on overhead projector transparencies if the teacher feels comfortable that way, gets more things done and can direct their energies towards teaching than suffer from teaching constantly interrupted due to the teacher’s poor skills in using technologies. There is, and can be, a gap between generations.*

6. *I don’t personally like to use tablet computers in class as I find contact teaching more natural. And I have never been too keen on using Moodle. But taking an advantage of technologies is important when studying, especially in the modern"
world. In the globalized world, most communication takes place through technologies so students should master these early to get the maximum benefit.” For instance, group work with strangers is much easier using the social media. For example, Facebook group discussions are a great tool.

The hybrid learners are, nevertheless, also aware of both the advantages and disadvantages of technologies in the classroom, and might even have a rather critical attitude towards their active integration in classroom activities. Therefore, these learners still appreciate the teacher’s traditional role as an expert in providing valuable information and being a stimulator for the learners. According to the hybrid learners, the use of social media should be voluntary in formal education.

In-school learners have always favoured traditional classroom learning and methods although in their out-of-school activities they have also relied on technologies for social relationships. These learners prefer concrete interaction with professional teachers, manuals, and paperwork in their learning process, although some use of technologies does not harm this process. See quotations 7, 8 and 9.

9. I would not allow students use any technological devices in class as they disturb concentration, are harmful for hand and eye coordination skills, and makes it more difficult to maintain discipline in class, which further disturbs peace in the classroom. Constant staring at virtual screens overstimulates the brain and can hurt the eyes.

10. The use of technologies is useful and saves time as such, but I think it has replaced contact teaching and real interaction too often. Challenging things are especially difficult to learn online. You need a qualified teacher and contact teaching. In addition, if the student has low motivation, online practice is easier to skip than a lecture or an exercise group. Contact teaching with good teachers is more motivating than studying alone at home in one’s spare time.

11. As a learner, I prefer traditional methods, taking notes and writing on the board, when teachers and students interact constantly. I do not believe that the internet or new technical devices facilitate language learning.

Despite their generally positive attitudes towards technology, the in-school learners have always felt they learn foreign languages better in classroom contexts.

5. Conclusion

This study examined extramural language learning, learning in hybrid contexts, and digital learning profiles. We focused on beginning university learners and analyzed their earlier experiences and behavior. According to the results, the participants were very familiar with extramural language learning. Its role seemed to strengthen as the learners’ proficiency and self-confidence grew. In particular, TV, films, and music often offered extramural language activities. The contexts of extramural language use diversified during education and also the foreign languages used multiplied.

Secondly, we analyzed which digital technologies were used and how the participants had used them in extramural language learning. The results suggested that the use of the social media and other spare time activities were most popular. Digital skills were learnt both in and outside school. Schools seem to play a role in introducing and providing support for the use of digital technologies. Mostly the skills needed for technologies were acquired outside school, often with help from family members and friends. In both extramural language contexts and technology use, the learner’s own engagement was found more important with age: the awareness of language learning processes and a desire to invest in the processes seem to increase.

Thirdly, despite the fact that the participants were so-called digital natives, they did not always associate digital skills with language learning in their reflections. Extramural language practice seems to be common, but the activities differ in number and variety from the ones used at school (see also, e.g. Ilomäki et al., 2012; Laakkonen, 2015).
When it comes to attitudes, technologies were considered helpful and good additions to traditional teaching and learning methods if teachers and users know how to use them (e.g., Benini & Murray, 2014; Mutta, Pelttari, Lintunen, et al., 2014; Thomas, 2011; Williams et al., 2014; Jalkanen, 2015). In our sample, attitudes towards technologies were mostly positive, but also critical.

Finally, we examined the digital learning profiles discovered in the learners’ answers. Three rough categories were identified to represent the main trends: digigage learners (heavy users of especially the social media, but who have not always mixed it with learning), hybrid learners (have used technologies, but with a critical mindset, for in and out-of-school learning) and in-school learners (have used technologies, but do not believe that they facilitate the learning process). The categories reveal variation in the digital generation: some learners do not consciously use their full digital potential to support language learning and try to keep certain areas of their lives less digital.

As our approach did not focus on individual longitudinal analysis, digital learning profiles could be analyzed further with quantitative (e.g., correlations) and qualitative (e.g., narratives, portfolios) data collection methods in future studies. A systematic longitudinal approach from primary to secondary education and beyond is needed to examine digital profiles and learners’ paths. This would reveal whether less successful language learners follow different paths than the highly motivated learners in this study or if digital learning profiles depend, for instance, on the learner’s age or target language proficiency.

The results corroborate earlier studies as learners seem to keep in and out-of-school activities separate, especially the use of technologies (e.g., Keltikangas-Järvinen, 2015). Changes in the flexible use of technologies might also be so rapid that the more rigid school system falls behind. This also means that surveys represent past situations, and even in classroom observation studies, the situation may have changed between the observation and the publishing of the results. One limitation of the study was the definition of technologies that can be understood in various ways as our results also revealed. In further studies, a more restrictive approach is needed to allow larger comparisons with earlier studies (cf. Healey, 2016; Thomas & Peterson, 2014; Wang & Vásquez, 2012). We chose not to direct or limit the participants’ responses too much in our questionnaires to survey their opinions.

In conclusion, foreign language learners are surrounded by hybrid learning contexts and often seem to be aware of various extramural language learning possibilities and know how to create affordances to promote their own learning. However, the digital age does not necessarily cause a linear addition of technologies in language teaching and learning: partly consciously, the potential of some technologies is not fully used to enhance foreign language learning. In and out-of-school use of technologies seem to develop at different, yet partly intertwined rates. Identifying learners’ digital profiles can support differentiation in foreign language teaching in hybrid contexts. Individuals and learning styles differ: some digigage learners create more actively personalized digital learning environments, whereas some prefer and rely more on classroom activities. Learners’ profiles could help teachers create suitable exercises or suggest new ideas to facilitate extramural and/or hybrid learning and assist learners in understanding their own learning styles and how to develop them in traditional and digital ways.

References


**Notes**

[1] “The Programme for International Student Assessment (PISA) is a triennial international survey which aims to evaluate education systems worldwide by testing the skills and knowledge of 15-year-old students”. Retrieved from http://www.oecd.org/pisa.


Reflective practice

The perceptions of a situated learning experience mediated by novice teachers’ autonomy

Paul Booth*
School of Humanities, Faculty of Arts and Social Sciences
Kingston University, UK

Isabelle Guinmard**
Institut des Sciences et des Pratiques d’Éducation et de Formation (ISPEF)
Université Lumière Lyon 2, France

Elizabeth Lloyd***
School of Education, Faculty of Health, Social Care and Education
Kingston University, UK

*P.Booth@kingston.ac.uk | **isabelle.guinamard@univ-lyon2.fr | ***Ea.Lloyd@kingston.ac.uk

Abstract
With the development of online language learning comes a growing need for courses in language teaching to incorporate educational technologies into course content. The challenge this development poses is how to incorporate educational technologies in teacher education programmes to prepare teachers for online language teaching. This study explores the way in which an authentic environment of English online and at a distance is facilitated by novice teachers and how their perceptions of the experience influence their own autonomy. The article presents how novice teachers cope with the complexity of the design of online materials, their pedagogy and their expectations. Data were collected via semi-structured interviews and novice teachers’ own evaluations of the course. The study found the opportunities and challenges for novice teachers in materials design, more complex roles and course expectations as they self-direct themselves in terms of both their learning and pedagogical skills. These findings suggest that teachers’ perceptions of situated learning can be shaped by their own teacher autonomy.

Keywords: Situated learning, autonomy, teacher education.

1. Introduction
The context of this study is a master’s degree programme based at a university in Greater London in which mostly novice teachers design and put online English language learning tasks for French learners of English based at a university in Lyon, France. The study was designed to enable language teacher educators to get a better understanding of novice teachers as they facilitate language learning in an online environment and how situated learning impacts on teachers’ autonomy. The aim is to put the novice teachers in a situated language teaching environment and for them to reflect on their experiences to develop their own schemas of pedagogy from authentic teaching. The significance of this study is that insights into how novice teachers grapple with facilitating CALL for their students in France are analysed in light of language teacher autonomy (Smith & Erdoğan 2008) of professional practice and cognition.
Situated learning is well suited to teacher education as authentic examples help novice teachers to reflect on their practice (Korthagen, 2010). Reflective analysis has also been used in online teaching for novice teachers to understand the core competences when teaching language at a distance (Guichon, 2009). This study argues how novice teachers’ perceptions of online CALL are shaped by their own degree of teacher autonomy.

Data for this study was collected from the trainee teachers’ own self-evaluations of their effectiveness as online facilitators. Semi-structured interviews were also obtained to corroborate data from the evaluations. This study analyses teachers’ perceptions of themselves, their tasks and their online language teaching experience in light of their own teacher autonomy. The perceptions help us to understand how novice teachers readjust to online language teaching and how this pedagogy is perceived.

2. Theoretical and empirical perspectives

Online language teaching has been well documented over several decades and in particular the challenges language learners face in this type of context e.g. O’Dowd & Ritter (2006), Lai, Zhao, Li (2008). The challenges novice language teachers face has also been extensively researched (Compton, 2009; Egbert, 2006; Guichon, 2009; Lai, Zhao & Li, 2008; McNell, 2013; O’Dowd, 2015; Slouti, 2007; Smith & Erdoğan, 2008; and Stickler & Hampel, 2015). However, Egbert (2006) argues that most language teacher education programmes work on the assumption that teachers can apply what they have learnt during the course training for CALL and then apply this knowledge in order to teach well. Egbert highlights that preparation of computer assisted language learning occurs in the confines of an education course. The problem is the lack of availability of extensive access to learners and authentic materials which means that many teachers find themselves unprepared for the challenges and realities of an instructional role (Egbert 2006, p.169). One way around this problem, Egbert suggests, is to situate the language teacher learning in CALL in an authentic problem-solving online context in which distance education gives novice teachers the opportunity to use educational technology with authentic language learners. Teachers who can analyse and understand the many different situations that may arise in technology-enhanced ESL classrooms will be more effective in helping their students to learn.

3. Situated Learning

Situated learning theory as Brown, Collins and Duguid (1989, p.41) explain, is when activity leads to perception and that both, moreover, are necessary precursors to the conceptualisation of ideas. This approach to teaching turns much traditional education upside-down in that conceptualisation; i.e., the formation of schemas needs to grow out of problem-solving activity. Knowledge, Brown et al. (1989) argue, is embedded in the situation and it is the circumstances that provide essential structure and meaning to learning. This approach to learning demands what Evans (2014) calls “a deep approach”; i.e., to see knowledge as complex, evolving, effortful, tentative and evidence-based (Evans, 2014, p. 187). Evans describes student teachers who manage to cross boundaries as those who transfer and adapt from what they have learnt from one context to another (Evans, 2014, p. 203). In our situated learning context, student teachers were required to transfer knowledge gained from the workshops on educational technologies for English language teaching to becoming facilitators in an online English language learning context for students based in Lyon.

The problem has been that trainee teachers’ comments tend not to refer back to instructional theory which highlights a difficulty of linking theory to practice in synchronous online teaching (Guichon, 2009, p. 181). The aim of this study is to put novice teachers in a situated language teaching environment for them to reflect on their experiences and to develop their own schemas of pedagogy which are analysed from a teacher autonomy perspective.
4. Situated learning in teacher education

Korthagen (2010, p. 104) argues for teacher education to encourage a pedagogy which combines experiences that help form the relevant gestalts. A gestalt is a combination of images, feelings, notions, values, needs or behavioural inclinations that form a whole (Korthagen 2010, p. 101). Korthagen (2010) explains that early teaching experiences tend to trigger gestalts in student teachers which tend to be related to classroom-based survival skills. What is needed in teacher education programmes is the further development of gestalts through suitable experience and subsequent reflection which is a process of schematization which the teacher educators wish to develop. This schematization would be a network of concepts, characteristics, and principles which are pertinent to the student teachers’ needs (Korthagen, 2010, p. 104). Our study uses novice teachers’ perceptions of their experiences after they have had time to reflect on their practice.

Huang, Lubin & Ge’s (2011) qualitative study compared a situated learning environment in an educational technology course with a traditional learning environment. Their study indicated that some pre-service teachers in the situated learning environment would prefer to be told what to do rather than to explore what to do on their own (Huang et al. pp. 2011-1209). Some of the participants appreciated the messy nature of their tasks and their own autonomy whilst others felt uncomfortable and frustrated (Huang et al. pp. 2011-1210). This dichotomy chimes with our own study in that some of the novice teachers seemed to flourish in an authentic learning environment whilst others wanted to be told what to do in certain situations. The degree to which pre-service teachers are autonomous may be an important factor in how they perceive situated learning. Huang et al. (2011) warn that educators must pay attention to pre-service teachers’ feelings.

The role of the instructor is important as well. Egbert (2006, p.176) highlighted that the role of the instructor was crucial in a situated language teacher web-based course on CALL which focuses on technology to support student language learning. In particular, Egbert highlights the need for instructors to know when to intervene to help student teachers and when to ask for authentic examples to enhance the situated learning experience. Egbert warns that there is the danger that student teachers can flounder in a mass of technological tools and so the instructor needs to scaffold the student teachers in order for student teachers to construct knowledge for real teaching contexts.

Situated language teaching in an online environment highlights the complexity teachers face in synchronous online tutoring. Guichon’s (2009, p.180) study of a teacher training programme for postgraduate students teaching French via a desktop videoconferencing platform showed that there were differences in participants’ accounts of critical episodes. Some participants merely described them whilst others identified the difficulty and examined their own pedagogy in a critical manner. So how novice teachers critically reflect is important to understanding their perceptions. O’Dowd’s (2015) study of telecollaborative exchanges also highlighted the need for information on the educational culture of the different countries involved.

5. Teacher Autonomy

How teachers reflect on their situated learning experience could be related to their teacher-learner autonomy. Autonomy is defined as the feeling of volition and choice in contrast to feeling pressured or coerced into action (Lynch, 2013, p.302). Smith & Erdoğan (2008) classify dimensions of teacher autonomy as two broad dimensions: professional action and professional development. In other words, as a bi-polar dimension: practice and cognition. The idea that practice and cognition are differentiated is important. As Borg (2009, p.166) explains, there is a strong relationship between practice and cognition in language teaching. Professional action relates to ‘the doing’ and as having three domains: the self, the capacity and the freedom (e.g. from one’s institution). Professional development relates to ‘the cognition’
and as having three domains: the self, the capacity and the freedom. Smith & Erdoğan (2008, p. 87) argue that what needs to be developed in teacher education is how to develop teachers’ capacity to self-direct their own learning. If novice teachers lack autonomy in their learning then they may lack autonomy in their teaching. In a language teacher education setting, Xu (2015) looked at teacher autonomy in collaborative lesson preparation and found that it is the type of collaboration which has a greater impact on novice teachers’ autonomy. Novice teachers’ professional development depends on synergies between their level of anxiety and the type of collaboration, therefore they should be given “more concrete help to scaffold their initial development, but not be offered ready-made resources to be directly used in their teaching. It is better for them to feel supported without being overly anxious and at the same time moderately challenged so as to promote their autonomy” (Xu 2015, pp 146–7).

The overall aim of the current study was to investigate novice teachers’ evaluations of their situated learning experience of online language teaching in relation to their teacher-learner autonomy. The emphasis on teacher-learner autonomy highlights the need to provide a rounded picture of the impact of teacher-learner autonomy on situated learning and its importance in designing teacher education courses in educational technologies for language learning and teaching. The study was designed to explore perspectives and experiences of novice teachers facilitating online CALL and the impact on their learning, pedagogy and freedom in course design.

6. Method

6.1. Background

Links were formed with a university in Lyon, France, and it was agreed with the tutor in France that the Lyon students studying for a Master’s degree in pedagogy would be recruited to complete the online language tasks and provide feedback to the London-based novice teachers on those tasks via a questionnaire. The Lyon-based students were all enrolled in the MEEF (1) course (Master in Teaching, Education and Training). From 2010, a Master’s degree is required in France in order to qualify as a primary school teacher. Students not only have to obtain this postgraduate degree but they also have to pass the national competitive exam allowing them to be employed. As the French primary school curriculum includes taking a foreign language, we decided to focus on the learning of foreign languages. Students cannot obtain their master’s degree unless they reach a B2 level in English (or any other foreign language). Moreover, most French students of English are not confident about their proficiency in English so it is challenging for them to initiate French primary children in the learning of English. This is why the ISPEF (2) Teacher Education Institute decided to set up a partnership with the university in London.

6.2. Participants

A total of 19 postgraduate novice teachers studying at a university in Greater London, were recruited from a module in educational technologies for language teaching which lasted one semester. The participants on this course had a range of L1 backgrounds: nine had English as a second or third language, whilst ten were native speakers of English. Thirteen had some sort of language teaching experience ranging from one-to-one to teaching groups. Prior to the course, online journals showed that their use of educational technologies included programmes such as PowerPoint, MSWord, DVDs, e-Podiums, interactive whiteboards and all of them had experience in using computers.

The MA module was comprised of eleven face-to-face 3-hour workshops. The workshops were highly interactive and were facilitated by two tutors who are both authors of the study. The topics covered on the course were: needs analysis and language level, planning tasks for language learning, distance learning and educational technology, designing tasks for online workshops, distance learning and student and teacher perspectives, peer and tutor feedback of students’ online tasks, reflection on own online
teaching, theoretical frameworks for evaluating teaching, learning and autonomy. During the course, the students published their own tasks on the Internet for the French students to complete within a period of four weeks.

6.3. The Virtual Learning Environment

The online language learning was delivered through the London university’s virtual learning environment (VLE). The system hosts teaching and learning materials for all of its courses. Students can download PowerPoint presentations, Word documents and PDF files, as well as contribute to discussion forums. For this particular course, a workshop platform (space on the university VLE) was used for participants to upload their online workshops. These mainly consisted of newly discovered technology artefacts like links to YouTube clips, images, interactive games, and video recordings and Word files that allowed them to design tasks to achieve specific language learning outcomes.

The London-based participants were given the role of online language facilitators. They were responsible for the content and the facilitation of language learning. The online workshops were constructed as part of the module assignment and the London-based tutors were there to give help and guidance to the MA students during the design and facilitation stages. The UK module workshops took place face-to-face once a week during one semester. The Lyon-based tutor was in direct face-to-face contact with the Lyon students to explain the rationale for the online course and actually took part in the course herself to gain first-hand knowledge of what was involved. Both the Lyon tutor and London-based tutors were in contact with each other via email.

6.4. Products from the learning environment

Each of the 19 UK-based novice teachers created an online distance workshop comprising of three online language learning tasks designed to last one hour. Their language groups consisted of either two or three French students who were grouped by English language ability using the Common European Framework Reference for Languages (CEFRL). Their language level was judged by their English teacher at their home university. The French students and UK participants were matched by the more experienced UK-based novice teachers facilitating the more advanced language students. The French students were encouraged by their French tutor to complete the
tasks outside of class but they were not obliged to do so. Below is an example a workshop for the French students.

Figure 2. The initial task to orientate the French learners to job interviews.

Figure 3. A vocabulary building task.
Figure 4. The next task focuses on the grammar of letter writing or job hunting.

Figure 5. A job advertisement for the learners to apply to.
The UK-based facilitators had access to a range of technologies and were encouraged to use multimodal teaching tools, including the use of video, image, sound and discussion forums. They were not required to create multimedia objects but were encouraged to find and utilise appropriate existing resources like YouTube and other video repositories, image databases, games and simulations. Some did create their own video or used Skype or social networking sites to communicate with the French learners but this was not a requirement.

6.5. Data sources

The data sources were designed to obtain insights into the novice teachers’ self-assessment of their own pedagogical and technological skills, and the situated learning environment in the context of teacher-learner autonomy. Following a qualitative research design, data was collected through two sources.

The first source was the novice teachers’ written self-assessment in acting as online language facilitators during a four-week period. They had to assess themselves according to five criteria: (i) the effectiveness of the tasks, including the use of technology in achieving the learning objectives; (ii) the quality of the teaching and learning taking place in the online workshop; (iii) the pedagogical choices made and how these impacted on learning; (iv) the challenges and limitations of facilitating a workshop in the online distance learning environment; (v) the implications for future practice.

The second source consisted of the semi-structured interviews conducted with five of the UK-based participants. The interviews were conducted at the end of the course and were audio-recorded and transcribed for analysis. The participants who were interviewed were reassured that their responses to the questions would not affect their final grades. Some example interview questions were:

- When you understood what was required of you for this module, what were your expectations of the workshops?
- What impact did this online learning and teaching experience have on you as a teacher?
What do you think are the advantages and limitations of online courses – both from pedagogical and technical perspectives? What do you think are the advantages and limitations of what we have done in the module? How will this experience change your professional outlook?

The actual tasks which the student teachers produced helped to show us how they developed language learning materials. The interviews were conducted by the researchers, who were also the tutors of this course, which helped us gain insights into the UK-based participants’ perspectives.

In order for the novice teachers to gain an understanding of how they facilitated online language learning, each novice teacher sent out an online questionnaire to her/his students in France. The questionnaire was written in both languages and the responses were gained from 35 Lyon students via five open-ended questions (see Appendix below).

6.6. Data analysis

The novice teachers’ evaluations of themselves and the interviews were analysed according to Thomas’ (2009) constant comparative method whereby evaluations are repeatedly compared for themes to emerge, which became the basis for analysis of the data. In addition, all of the data from the recorded interviews were transcribed and analysed for themes. The themes from the evaluations and interview transcriptions were combined and connections were drawn between them to create three overarching themes in relation to teacher-learner autonomy.

7. Results

The thematic analysis revealed three superordinate themes: self-directed learning as a teacher, which focused on areas such as creativity and task design; teacher-learner experience in relation to professional development, which focused on areas such as learner and teacher autonomy, time-management, experience, roles and misunderstandings; freedom to self-direct one’s teaching, which focused on areas such as expectations and complexity. Each of these will be discussed in turn.

7.1. Self-directed learning as a teacher

Creativity. The experience of authentic online teaching gave some of the participants a sense of freedom for them to experiment with different technology in order to foster language learning. There were creative uses of YouTube, for example, in the workshops that gave the participants the inspiration to experiment and to develop tasks which they had never done beforehand in an online environment. Participant 7: “So it [online tools] definitely helped me in that sense, being creative and being innovative with what I can use in the classroom”. Here, the online setting is felt to offer novice teachers more opportunities to use technology in innovative ways and that creative use of computers demands creativity and commitment from the facilitator, as participant 14 commented: “I think that making use of computers to enhance learning is also related to the creativity and commitment of the facilitator with his or her teaching career”. Here we can see that the more self-directed novice teachers embrace technology because it offers a way for them to express their creativity and that they are not daunted by their lack of experience.

Task design: freedom and boredom. Along with the perceived creativity came the idea of freedom to make their own choices based, partly, on a needs analysis which they were encouraged to produce and send to their students in Lyon.

From a teacher’s perspective, an advantage is the level of freedom even though I had to stick with the needs analysis questionnaire, that still gave me a variety of things to choose from and to work towards and I think even creating the materials it helped me improve my skills as a teacher cause not only the pedagogical side to it, unless materials development is the pedagogical side, but I feel like that really
helped me to think of different ways... to think how to use a simple thing like movie and base my lesson on that or at least make it the beginning of the lesson. So I think it pushed me as a teacher and there was pressure but I feel like it helped me develop as a teacher. (Participant 11)

This participant is starting to think about materials development and how this process makes her reflect more deeply on pedagogy. Materials design for this participant could be the springboard to schematise her own pedagogy.

Whether you are online or in a classroom I feel like I’m more analytical in my work. I’m analysing it more in terms of thinking of the students more. Before, I used to plan lessons and it was like you know ...but now you need to consider so many things, especially because you introduced the concept of needs analysis. Because normally when you’re sent to a school or college or university you go straight into a classroom so you don’t need to do a needs analysis because it’s already been done. But this really made it personal. And it really made you think critically about the way you plan lessons and staging everything. (Participant 12)

This participant starts to be more analytical. Moreover, she is starting to understand the greater responsibilities placed upon her as she can no longer expect to be given a syllabus as part of a language learning programme. She also expressed the need for more technological variety in order to encourage her learners to return to the tasks each week.

However, along with the freedom, participant 2 started to find negative consequences in the design of his own tasks.

However, as I was new to adapting materials on a digital platform, I designed tasks that may not have been appropriate in terms of linguistics and usability. However, such practice exposed me to the idea that there is always a possibility that educators will encounter all sorts of difficulties and that they should be ready to find alternative solutions. While I was introduced to the exciting possibilities of online teaching, I also became all too familiar with the challenges both teachers and students face in a digital age. (Participant 2)

This participant sees the challenges of materials design and their suitability. For example, in the feedback from one of his students, the tasks were perceived as being the same as pen and paper exercises. This negative feedback highlights how for one of the novice teachers he did not experiment with different ways of using online technology and did not design meaningful tasks.

7.2. Teacher-learner experience: professional action

Learner and teacher autonomy. Two of the participants interviewed expressed the idea that online learning encourages learner autonomy.

It makes it easier for the learner because there is a bit more autonomy for them because they can kind of go at their own pace or they can get in contact with you when they want to without being pushed like they are in a classroom. (Participant 7)

Likewise, participant 12 puts forward the idea that learner autonomy could feed into teacher autonomy.

It definitely encourages learning autonomy. By doing online learning, distant learning you are learning about your own strategies of learning and I think that helps a lot, because I’ve learnt a lot about myself... learning to teach online. Just my time management, in terms of sending out emails, communicating with my students... am I thinking for myself or do I get help. Little things like peer support; you know getting my friends to help me do something or figuring it out. I found that I actually prefer to figure things out myself to understand it. Or just asking my friend teach me, I’d rather just do it myself because I feel that’s the only way
you are going to learn. So I felt that had an effect on the learners as well. They were doing the same thing. Because there were a lot of technical problems and that’s one of the disadvantages. (Participant 12)

This teacher starts to reject peer support in favour of relying on her own ways of dealing with students and starts to become more self-sufficient and understand that learning starts from the self. However, participant 12 who had taught English in Saudi Arabia could draw upon her own experience in solving problems which were usually technical in nature; e.g. Lyon students not being able to log on to the VLE.

**Lack of experience.** One participant was particularly at a disadvantage as she lacked experience preparing language courses. This course had a negative impact on her as she felt that she was under-prepared to take on the role of an online language learning facilitator.

> Given this experience is the first lesson I have created as a teacher, by doing this workshop it really challenged my job as a teacher in the future and hinted the areas I need to immensely improve for my career as a successful teacher in the future. This task was a big eye-opener for me and taught me a completely different aspect of a role of a teacher I never have been exposed to back home. (Participant 5)

Her thoughts highlight the need for language teacher educators to take into account the individual support that some participants may need compared to others. Clearly if novice teachers do not feel as though they have enough experience to cope with a situated learning environment then this approach can have negative consequences. If, however, novice teachers perceive themselves as having the capabilities to learn in this type of environment then this type of experience will help them to flourish.

**Time-management.** The pedagogy of online language teaching highlighted the amount of time the online workshops took to prepare. Time-management was highlighted by two participants (13 and 18) as an important skill to develop. Coupled with the fact that these participants did not have experience of online teaching, meant that they did not have a repertoire of tasks to draw upon. The time-consuming nature of task preparation made them realise that they needed to schedule in more time for task preparation than perhaps anticipated.

**Complexity of online management.** The time-consuming nature of task design, the issue of “not being in ‘direct’ supervision by the teachers, and the freedom to carry out tasks independently” (participant 14) increases the probability of misunderstandings in communication between the language learners and the novice teachers. This quote shows how some of the novice teachers felt under pressure. The complexity of online management was also commented upon by another participant who saw that her role needed to be highlighted to a particular learner.

> I was a motivator also in that I encouraged the participants to regard the workshop as a platform for improving their L2 proficiency and not to prove this. For instance, the A2 participant was nervous as she tried to produce the language ‘perfectly’ although she was a false beginner. Assurance was a necessary psychological calming tool in assisting her to produce language. (Participant 3)

This novice teacher starts to take on the role of calming the student; to think of the tasks as improving rather than perfecting language.

7.3. **Freedom to self-direct one’s teaching**

**Expectations.** Throughout the online workshops, novice teachers had the freedom to self-direct their own teaching. This was seen by most of the participants as positive in that teacher action was not constrained by a pre-set syllabus and that novice teachers could develop their own teaching practice. However, there was another side to this freedom in that it was seen by one participant as indicating lack of direction.
And I think that got a tiny bit frustrating because obviously there was times when tutor 1 and tutor 2 were not in the same room at the same time so she [tutor 1] would say something and you [tutor 2] would say something different and we were like which one shall we do, which direction shall we go in? Because again this is completely new for us and I really sympathise with the students that I’ve never taught before. Those that have a teaching background know how to pick up on these problems. But some of us who have never taught, I don’t know what they were going through. They were completely puzzled at some point. I don’t know. (Participant 12)

This participant seems to expect to be told what to do when a problem arises in a situated learning experience, which she felt to be frustrating. Some of the novice teachers may have wanted definite answers to their problems when faced with online teaching. This particular participant did not appreciate the different perspectives of her tutors, which she thought would confuse students with little or no experience. There was also confusion in terms of expectations from the students in Lyon as some of them were not apparently aware of the goals of the online workshops.

I think a lack of knowing the situation in terms of we were told that the students had to do this course but after speaking to the students I felt like they were as confused as we were in terms of how important this English course actually was. Because I kept getting questions like ‘I don’t understand why I’m doing this’, ‘what’s going on?’ Things like that so that made it difficult for me because I thought I had to deal with the administration and learning how to do this thing that I’d never done before so that was one thing. So I felt that maybe there was a miscommunication between the significance of this course for them. And that had an effect on us giving them these lessons. (Participant 12)

Clearly, miscommunication was perceived regarding why the online course was taking place and the teacher felt that it was not her role to explain the rationale for the course to her students. She thought that her role was simply to design tasks and facilitate online language learning. In a situated learning environment novice teachers may not comprehend the complexities involved until the course is running. When problems occur, as they clearly did, some of the novice teachers immediately turn to the tutors and expect the ‘right’ answer to solve their problems. If they do not get a straightforward answer to a complex problem then they may criticise the pedagogical approach and misunderstand the greater demands placed on them in terms of responsibilities.

8. Discussion

8.1. Teaching experience and reflection

Several themes have emerged from this study which can inform teacher educators. The study revealed that when given responsibility, novice teachers react in different ways. Teachers with less experience sometimes tend to feel at a loss. This reaction is in line with Guichon (2009, p. 179) who argues that the capacity to infer from teaching practice might depend on former experience, amongst other variables. Those with more experience were able to reflect on their practice and offer insights into how they would change as a result of that experience. Therefore, although situated learning may present opportunities for critical reflection, teacher educators need to balance situated learning with the backgrounds of the trainees and provide more scaffolding for reflection to those with little or no experience.

8.2. Greater responsibility

The price paid for the partnership between the two institutions translated into greater complexity in terms of there being more responsibility for the novice teachers and their learners. This complexity tended to be more acutely felt by the facilitators who had relatively little teaching experience. Some of the novice teachers were not comfortable
having to solve problems by themselves. When they asked their tutors for help and advice on how to deal with lack of cooperation from their online students, they were perplexed when they did not get straightforward answers to their problems. However, complexity also initiated peer-supported learning outside of class.

The London-based teachers described themselves as having a greater degree of autonomy than they had perhaps previously experienced on other courses. Autonomy was expressed both for themselves as facilitators and for the language learners in France. Critical evaluations from trainees who reflected on what they could have done differently in the workshop to improve communication and learning tended to show more autonomy, whereas the trainees who tended to highlight the limitations of the technology rather than reflecting on their own practice, feeling powerless to affect change in this teaching environment, showed less autonomy.

Banegas and Busleimán (2014) found that autonomy, i.e. studying alone, managing deadlines and one’s own learning pace had the biggest impact on motivation in an online pre-service language teacher education programme. This whole aspect of self-management can be seen as empowering for novice teachers who wish to pursue a career in teaching as it gives a sense of freedom but it comes with responsibility. In this current study there was a need for the Lyon students to participate as the London-based teachers had an assignment to write which would depend on the Lyon students’ engagement with the course. This dual role which the London-based participants experienced as both learners on a master’s course and online facilitators is essential for a deep approach to learning to teach (Evans, 2014).

8.3. Schematisation

One aspect that teachers reflected upon was their freedom to design their own language learning materials. This shows that teachers are starting to schematise, i.e. link theory to practice. Tomlinson (2013, p. 482) supports the notion that in order to cater for students’ needs in language learning, teachers need to develop language learning materials and the process of creating these materials helps teachers to engage with the ideas and theories of second language acquisition. Moreover, McNeil’s (2013) study showed the importance of the freedom for teachers to create activities of their own for their learners. Therefore, materials design could be one of the triggers to help schematisation of novice teachers.

There are several ways to aid teacher autonomy in a professional environment. One way, as we have seen, is for novice teachers to write reflective journals on what they have experienced. Smith & Erdoğan (2008, p. 89) highlight this approach; however, they also highlight the danger that if these journals are assessed, novice teachers may only be writing what they think their tutors want to read. The processes that the assignment encourages may be more important in terms of fostering deep learning than the product itself (Gibbs and Simpson, 2004-5, p.15). Another theme which became important for teacher autonomy was the freedom to design their language learning tasks. Materials design encourages teachers to reflect on second language acquisition and, thus, develop their own schemas of knowledge. Finally, the results have shown that teachers need the freedom of the institution to solve problems for themselves. This can develop deep learning but may come at a cost as not all novice teachers thrive in this type of situation and so may need extra help and scaffolding techniques.

9. Conclusions and implications

A situated context throws into perspective how important it is for both facilitators and learners to be engaged in the course from the beginning. From a pedagogical perspective, this means novice teachers become acutely aware of their perceived lack of experience and their lack of technological understanding. The more autonomous facilitators found that they engaged more profoundly with their own understanding of language teaching. In this respect, situated learning was beneficial. When a facilitator was not able to draw upon theory and reflect upon his or her own schemas of
pedagogy, the frustration tended to be directed toward dissatisfaction with the course or the technology. The cost of situated learning was the increased complexity of the language facilitation for language learning.

One possible solution to this difficulty is to engage novice teachers in greater reflection. Although reflection through learning diaries is well established, reflection is not entirely obvious to all students. A reflective cycle (e.g. Dewey 1910 in Roberts, 1998) model can be made explicit which focuses attention on 'critical moments'. These crucial moments, which can arise from problems, may focus reflection. Subsequently, problems are seen as inevitable and an opportunity to move from the factual and apportioning of blame to understanding the problem as a pathway into theoretical explanations. As Teasdale (1993) argued, deeper understanding is gained not through a focus on the factual but through implicit understanding of our relationship with factual knowledge.

Online teachers and learners have deep-rooted expectations about the nature of the course. From the facilitators’ point of view, explicit explanations regarding what the course entails is necessary. One of the fears novice teachers have is their lack of proficiency with technology. In fact, no expert knowledge of technology was required and yet this was highlighted by one interviewee. The point is that there may be all manner of misunderstandings which can vary from cohort to cohort. However, explicit and upfront explanations of what the course is and is not can help dispel fears. As the online course is supplemented by a face-to-face course in France, there needs to be greater co-ordination between the two institutions. In other words, better links need to be made between online language learning and face-to-face teaching and learning.

This study sheds light on how important the social aspect is for engagement and for communication, which is in line with Compton’s (2009) community building and socialisation strategies. Without meaningful interaction between student and facilitator, the tasks may seem like any other exercise. When there is a genuine desire or need to communicate then student output is ‘pushed’ (Swain 1995). This communication is essential for second language development. When communication is formulaic or when the French learners do not feel that they need to express themselves then engagement tended to drift away. A possible solution to develop links between the two cohorts is to give a personal presentation at the start. With subsequent groups of learners and facilitators, a Skype or recorded presentation was made by the London group of facilitators and some of the Lyon students reciprocated either by email or by Skype.

On an assessment level, the French students need marks to be evaluated for the English course. There needs to be a balance between formative feedback from the facilitators and summative assessment for the French education system. We came to understand how important the possible washback effect is from assessment and how appropriate assessment strategies need to be fully addressed in building professional skills in teaching and language learning. This is of course a fundamental reason for creating a situated learning environment and, despite the difficulties outlined above, our students generally spoke positively about their experience in this module.

Notes

Acknowledgements
We are grateful to all the Masters’ students who participated in this study. In particular, we appreciate the help from Julia Deschamps who contributed to making this project possible.
References


**Appendix: The questionnaire given to the Lyon based students**

1. L’atelier atteint les objectifs pédagogiques? Expliquez votre réponse.
   [Does the workshop achieve the stated learning objectives?]

2. L’atelier faciliter l’apprentissage de la langue anglaise?
   [Does the workshop facilitate second language acquisition?]

3. Les matériaux multimédia sont utilisés efficacement pour aider à améliorer et consolider l’apprentissage de la langue ?
   [Are the multimedia materials used effectively to foster and consolidate language learning?]

4. Les matériaux engagent l’apprenant?
   [Do the materials engage the learner?]

5. Les matériaux pédagogiques donnent-ils l’opportunité à l’étudiant d’apprendre l’anglais par un processus cognitif profond ?
   [Do the materials encourage deep cognitive processing of the learner?]

Answers to these questions were not included in the analysis but fed into the novice teachers’ own evaluations.
Editorial Board

Dr. Christine Appel (Universitat Oberta de Catalunya, ES) | Dr. Paul Bangs (London City University, UK) | Dr. Françoise Blin (Dublin City University, IE) | Dr. Alex Boulton (Université de Lorraine, FR) | 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. Phil Hubbard (Stanford University, USA) | Dr. Kristi Jauregui (University of Utrecht, NL) | Dr. Mike Levy (University of Queensland, Australia) | Dr. Peter Liddell (University of Victoria, Canada) | Dr. Antonio Martínez-Sáez (Universidad Autónoma de Madrid, ES) | Dr. Robert O'Dowd (Universidad de León, ES) | Dr. Ana Oskoz (University of Maryland, USA) | Dr. Pascual Pérez-Paredes (Universidad de Murcia, ES) | Dr. David Perry (Universidad Politécnica de Valencia, ES) | Dr. Joan Tomàs Pujolà (Universitat de Barcelona, ES) | Dr. German Ruiperez (Universidad Nacional de Educación a Distancia, ES) | Dr. Bernd Rüschoff (University of Duisburg-Essen, DE) | Dr. Mathias Schulze (University of Waterloo, Canada) | Dr. Rafael Seiz-Ortiz (Universidad Politécnica de Valencia, ES) | Dr. Ana Sevilla-Pavón (Universitat de València, ES) | Dr. Bryan Smith (Arizona State University, USA) | Dr. Peppi Taalas (University of Jyväskylä, FI) | Mrs. June Thompson (University of Hull, UK) | Dr. Steven Thorne (Pennsylvania State University, USA) | Dr. George S. Ypsilandis (Aristotle University, GR)

The EUROCALL Review is 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](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.
- 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 Ana Gimeno, email agimeno@upvnet.upv.es

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