VISP: a MALL-based app using audio description techniques to improve B1 EFL students’ oral competence

ANA IBÁÑEZ MORENO (UNED), ANNA VERMEULEN (GHENT UNIVERSITY)

The last decade, mobile devices, defined as “any device that is small, autonomous and unobtrusive enough to accompany us in every moment” (Trifonova et al. 2004:3) are a familiar part of most teachers and students’ lives and their integration into learning and teaching is constantly increasing. In this contribution we concentrate on creating learning materials for mobile devices in order to suit the current ‘On Demand Generation’ that wants “anything, anywhere at anytime” (Caudron 2011: 25.). In order to explore the benefits of mobile learning away from the traditional classroom, we designed an app named VISP (VIdeos for SPeaking), which is available in Android operating systems. VISP consists of using audio description (AD) techniques to practice oral language skills. We present here the methodological steps followed in the design of VISP, as well as the first qualitative data obtained, derived from applying it to two Spanish students of English as a foreign language (EFL).

Keywords: MALL, app, mobile phones, EFL, didactic tool, audio description (AD)

1. Introduction

From the beginning of the eighties of the 20th century, audio-visual products such as films, theatre and media events began to be provided by means of oral descriptions in the gaps between dialogues in order to make them accessible to visually impaired people. Audio description (AD) as a mode of audio-visual translation (AVT) was born. Studies such as Schmeidler & Kirchner (2001), Whitehead (2005), Bourne & Jiménez Hurtado (2007), Snyder (2011), etc. have shown that this new mode of AVT not only increases the accessibility to audio-visual products for the visually challenged people, but also for children, elderly people and immigrants. Additionally, it has recently been applied in the foreign language (FL) classroom to improve the written skills of English as a foreign language (EFL) students (Clouet 2005), as well as the oral, lexical, phraseological, grammatical and even intercultural competences of Dutch-speaking students of Spanish as an FL (Ibáñez Moreno & Vermeulen 2012, Ibáñez Moreno & Vermeulen 2013, 2014).

In this line, we applied it to mobile learning (m-learning), and developed a mobile application (app) for EFL students to promote oral skills. We present here the first version of VISP (VIdeos for SPeaking), which contains a short clip (31 seconds) of the film Moulin Rouge (Luhrmann 2001) that has to be audio described. The process of audio describing a video clip using this app comprises several steps: 1) filling in an online pre-test/questionnaire, 2) viewing the clip several times, 3)

---

1 The research presented in this chapter has been written in the wide context of the SO-CALL-ME project, funded by the Spanish Ministry of Science and Innovation (ref.no.: FFI2011-29829).
drafting a small (written, if they wish) AD script (ADS), 4) recording the ADS and sending it, and 5) filling in a final test/questionnaire containing formal and functional questions related to the task, where users can assess themselves by comparing their AD to the original AD. This first prototype was tested with two distance education students of EFL, who were studying to reach level B2 (following the Common European Framework of Reference for Languages, CEFR, 2001). In what follows a brief state of the art regarding the fields of mobile assisted language learning (MALL) (point 2) and AD as a tool in the FL are presented (point 3); then, the data obtained from the qualitative pilot study are shown and analysed (point 4). The results are positive, although more research is suggested to assess the validity of this study (point 5).

2. MALL apps: state of the art

MALL has been defined by Kukulska-Hulme (2013: 2) from the perspective of the connection between language learning and the new mobile technologies: “mobile technologies in language learning, especially in situations where device portability offers specific advantages”. This concept is intimately linked to the development that mobile technology has experienced in the last decade as well as to the wide variety of emerging mobile devices. Despite the wide range of apps that help improve FL skills, if we focus on those that have been specifically developed within the academic world, there are not many. As Rodríguez Arancón et al. (2013), Calle-Martínez et al. (2014), Martín Monje et al. (in press), Arús-Hita et al. (in press), and Pomposo et al. (forthcoming) remark, there still a need for empirical studies based on testing the available apps yet, especially with regards to oral competences. Some pioneering work is being undertaken by the research group ATLAS (Applying Technology to LAnguageS²), whose members have designed MALL apps to promote oral skills by enhancing the communicative competences: ANT (Audio News Trainer, Pareja-Lora et al. 2013), FAN CLUB (Friends of Audiobook Network), VIOLIN (Videos for ListeniNg, Elorza et al. 2014), BUSINESS APP, EATING OUT, and MARLUC. Given that this field is still under development, our aim is to contribute to it, in line with ATLAS research group, by proposing a MALL app with strong academic grounds and pedagogically solid, as well as motivating and stimulating for EFL students.

3. Audio description (AD) and its application in the FL classroom.

The concept of providing oral descriptions in the gaps between the dialogues for visually impaired people was first developed in the US, in the 1970’s (Snyder 2005), and introduced in Europe at the beginning of the 1980’s (Benecke 2007). The so-

---

called AD turns the visual (the images) into verbal (speech) in order to make audiovisual material (films, theater, opera, etc.) accessible for the blind and partially-sighted. Like other modes of audiovisual texts—such as subtitling and revoicing—it is a non-autonomous form of linguistic transfer, since the lexical choices and the syntax must be connected to the audio and the image of the audiovisual material. Consequently, the coherence of an AD is based on semiotic cohesion: the interaction between the sound and the image.

Due to its social relevance and its multidisciplinary approach it has become more and more attractive for scholars, who have studied this kind of intersemiotic translation from different points of view. From the translating point of view, Matamala (2006) and Díaz Cintas (2007) among others have dealt with the competences and skills of the good audio describer, who has to be aware of his/her role as a social intermediary. Peskoe (2009), who explores the educational benefits of AD for both blind and sighted preschool children, shows that the depth of information provided with AD helps to develop more sophisticated language skills. That brings us to the use of AD as a didactic tool in FL learning, which is discussed below.

The multimodal nature of audiovisual texts, which combine the verbal sign with images and sound, represents a stimulating background for language learners. In line with the success of applying subtitling and revoicing to language learning and teaching (Sokoli 2006, Diaz Cintas 2008, Borghetti 2011, Incalcaterra McLoughlin & Lertola 2011, Talaván Zanón 2013, etc.), AD proves to be beneficial as a didactic tool in FL learning. Clouet (2005) first proposed the use of AD as a didactic tool to promote writing skills in English as a FL. Martínez Martínez (2009) conducted a successful experiment on the possibilities of self-learning vocabulary in FL through AD. Kreitz et al. (2012) also conclude that the extra narration proves to be successful in educational settings for sighted children. The EU-funded Project ClipFlair, which promotes multilingualism in Europe, also provides activities based on AVT to foster language FL learning, including AD-based ones. In the same vein, the ARDELE project at the University of Ghent (Belgium), explores the benefits and the limitations of the use of AD in the Spanish as an FL classroom as well as in translation courses. To date, the results on the analysis of the students’ learning outcomes show that AD enhances lexical and phraseological competences (Ibáñez & Vermeulen 2012, 2013, 2014).

---

3 ARDELE stands for Audiodescripción como Recurso Didáctico en la Enseñanza de Español como Lengua Extranjera

Once the validity of AD as a didactic tool in the teaching and learning of FL in formal settings, we have conducted a study, by designing an app and testing it. This app is presented below.

4.1. Preliminaries: theoretical framework and pedagogical premises

VISP (which stands for Videos for Speaking) is a mobile app that has been conceived within the framework of MALL for the promotion of oral production skills. The reason is that even if ubiquitous learning environments have increased and new technologies have been developed to adapt to the new learning styles (Jones and J. Ho 2004), there are fewer chances for the average user, in an average ubiquitous context, to practice oral production. This first version (VISP v1) has been conceived for students of EFL who possess at least a B1 or B2 level (according to the CEFR 2001, 2007) and who are interested in practicing their oral skills while improving their lexical competence, which is considered by many as the primacy aspect for Second Language Acquisition (Tight 2010). This app has been developed departing from a number of theoretical premises: first, from the language learning perspective, the importance of lexical and phraseological competence for proficiency levels (Tight 2010) is higher than other aspects of language learning. That is, lexical errors are perceived by native speakers as more serious than grammatical errors (Schmitt 2000). Moreover, phraseological competence is the most difficult competence to acquire, and it does not come implicitly –that is, naturally– (Laufer 1997: 25). The number of exposures necessary for both the recognition and the production of vocabulary is significantly lower than the number of exposures needed to learn words in context (Nation 2001: 289-299, in Tight 2010).

Second, from the perspective of mobile learning VISP v1 is conceived as an ubiquitous learning tool: personalized learning and learning based on user-generated contexts (Traxler 2011), to be used in those moments when the user has nothing else to do: when he/she is waiting for the bus, the metro or the train, or in the waiting room of an office. The user only needs a few minutes to accomplish the task. He/she can pause the application, and continue or repeat the activities whenever he/she wants. The activities of one session, altogether, take an average of 30 minutes, but the user may devote more or less time, depending on his/her time and motivation. VISP is currently compatible with Android. Illustration 1 shows the home screen of the application:
Thirdly, from a pedagogical perspective VISP v1 is framed within the currently accepted and used language teaching approach: the communicative approach. Within it, VISP makes use of the transferrable skills approach (Talaván Zanón 2010 and others), which is necessary in ubiquitous learning to support the idea of using mobile devices for language learning, and of the task-based approach (Willis & Willis 2007) in the sense that it consists of communicative tasks whose goal is to achieve a specific learning objective (Ellis 2003). In both approaches the user develops his communicative competence: when assuming the role of an audio describer, the learner must take the recipient into account in a way that helps him become aware of the fact that what he sees and interprets does not contaminate the objective idea that the recipient has to receive. In the next section VISP is briefly described.
4.2. Description of VISP v1.

For this pilot study, we selected a clip of the film *Moulin Rouge* (Luhrmann 2001), which lasts only 30 seconds, as opposed to the clips that have been used when working with AD in the classroom (Ibáñez Moreno and Vermeulen 2013, 2014), which last around three minutes. Following the guidelines (Benecke 2007, Matamala and Orero 2007) an audio describer is allowed to use 180 words per minute. In this particular case, the user will employ around 55 to 60 words. The clip was chosen because it does not contain dialogue, and not a lot of action happens (which gives the user time to describe what he/she sees). The ADS of the whole clip is provided below:

(2)

A handsome man, Christian, in his twenties, with dark hair and beard, takes a new line on his typewriter. He puts his hand to his forehead. Through his open window lies Paris at night. Tearfully, he stares out to the window, at the Moulin Rouge. He turns back to the typewriter. The Paris city scene.

Clips should also be selected on the basis of their narrative content, according to the vocabulary the user has to learn. In this case, the words and expressions that the user will learn are, mainly, the following: *a boy in his twenties, takes a line on his typewriter, puts his hand to his forehead, Paris lies at night, he stares out of the window, he turns back to his typewriter.*

Second, the instructions for the usability of VISP v1 were designed. They consisted of several steps, beginning by a very brief introduction to AD. The image below shows the *Introduction* screen:

---

4 The AD was commissioned by Fox International and provided by IADA Ltd.; The script was written by Patrick Mulcane and narrated by Dennis Lawson, produced by Carol McGregor of IADA Ltd. at Murricane & Murricane Studios in Glasgow, Scotland (IMDb).
As can be observed, only the most essential information about AD was included. This page has two buttons. The one above (Sample Audio Description) directs the user to a real AD sample: when the user clicks on that button, he gets access to a clip of 4 seconds, extracted from *Memoirs of a Geisha*, with AD.
In this short clip of 4 seconds the user listens to a real AD, as a warming-up listening task. Finally, the last button, at the bottom of the screen directs users to a Google Drive questionnaire, where they fill in their personal data: name, surname and email and completes a short test, which includes language content that will appear on the AD task and therefore on the ADS. Also, this questionnaire includes a link to the YouTube trailer of the movie from which the clip is obtained: Moulin Rouge. This video is also meant as a part of the warming-up phase, in the sense that users can start familiarizing themselves with the audiovisual material that they will encounter immediately afterwards for the activity.

Once users have been introduced to AD, seen and heard an example, and filled in their data, they can continue to the next step, by clicking in the main menu on the Instructions button, which takes them to a screen as the one shown below:
These instructions are very simple, brief and direct, so as to keep the user’s attention and interest. They contain tips and basic rules to audio describe. Once users read the info contained in this screen, and they know what they have to do, the next step is the task itself, as seen below:
In this screen users practice as much as he/she wants with audiodescribing, by recording themselves, listening to themselves, repeating, changing their AD. Once they are satisfied with their performance, they go back to the Home screen and, once there, click on the button Finish (see illustration 1 above). This button directs them to the screen in illustration 6 below:
In this screen users just fill in their name, and send their recording. The recording is sent to an e-mail account. In this way, track can be kept on all the users’ tasks, which are received in an mp3 attachment to their email, with their name and surname. Besides, this screen also includes a self-evaluation section, to which the user accesses by means of clicking on the button below the send button. This final questionnaire has been conceived to obtain data on the user’s perception of the application and on the idea of using AD as a learning tool, and, to help him/her assess his/her own performance. In order to do this, the first thing that users see is the original clip with AD, as well as the written original ADS. Thus, they can both listen to it while they read, and see how the audio describer talks during the gaps with no dialogue or off-voices. This is a first self-assessment step: watching the original ADS and seeing how it was done. In this step they can start comparing their own AD with the professional one. A second step consists of more specific self-assessment, where reflection on their performance is elicited by comparing, more specifically, their own ADS to the several parts of the original ADS. For this, several questions are included that elicit the user’s reflection on their own learning process, thus using meta-cognitive skills and eliciting awareness. They are outlined below:

2. Now, let’s compare it to your ADS:
a. Which words/expressions are different from yours?

b. In any case, any interesting differences on what you said?

c. Did you use expressions such as “We see, we observe…” in your ADS?

d. Did you use adjectives or adverbs in your ADS other than the ones of the original ADS?

(2.a) is a multiple-choice question. That is, it contains several options (all the parts of the original ADS are included), and the user has to tick the ones that he considers adequate: all the expressions that he used completely differently (that is, where not even three words are the same). We cannot expect learners to produce a totally similar ADS to the original, because many texts can be produced based on one same source text. Users may be discouraged if they are expected to produce an AD that is exactly the same as the AD in the clip. Therefore, we think that the fact that he/she has to use at least three similar words is an appropriate limit to evaluate his/her own performance. In a first pilot study, however, we aim to obtain data on the average of similarities and differences between users’ ADs and the original AD, which will allow us to adjust this number.

As regards (2.b), it is an open question, where users can write their particular observations on the differences between his AD and the original. This question is meant to leave some freedom and autonomy to the learner to analyze his own performance without instructions.

Finally, (2.c) and (2.d) are Yes/No questions. Below them, the following notes can be read, respectively:

(4)

2.c. The key issue here is: Do you think it is polite to remind a visually impaired person what we can see? Besides, we have limited time, and these expressions just steal it from us! Think about that :-)

2.d. As you may have noticed, in the original ADS there are indeed adverbial expressions such as "tearfully". Now: compare "tearfully" to saying "he looks sad". The adverb is just describing what appears on the screen, while the sentence is more subjective, isn't it?

These remarks are meant as guided reflection, to make users take their (supposed) audience into consideration: visually impaired persons. Therefore, these two questions are meant to elicit interpersonal and intercultural competence, in the sense that the audience is taken into account and the user is encouraged to realize how important it is to communicate a message adequately, and how important it is the message we communicate and how we communicate it.
Finally, section 3 in the post-questionnaire contains several grid questions, where users must evaluate (in a degree of 1 to 5, where 1 means “I totally disagree” and 5 means “I completely agree”) the use of AD for language learning purposes. This section contains four questions, reproduced below:

(5)

a. This AD task has made me think about my own language learning.

b. This AD task has made me reflect on how what we communicate is strongly influenced by our particular way of looking at things.

c. AD has helped me observe the importance and difficulty of using accurate vocabulary.

d. This task has been useful for me to observe how important it is to make all type of audiovisual material accessible to visually impaired persons, by using the language in such a way that the recipient is taken into account.

These questions are intended to make users reflect on AD as a technique to improve certain areas of their vocabulary in English, as well as on their own communicative competence in general. They are meant to assess: their general view of the benefits of this technique to improve language learning (6.a), their view on AD as a tool to make them think about the act communication (6.b), their view on AD as a tool to realize the importance of using accurate words (that is, to develop lexical competence awareness) (6.c), and their view on AD as a tool to grow awareness on others, in this case, on visual impaired people (and therefore, to develop interpersonal and intercultural competence) (6.d).

5. A qualitative pilot case study

In order to test VISP v1 a first pilot study was conducted during the second term of school year 2014-2015, with two students –who will be called student A and student B– from the Degree in Tourism studies of UNED (Universidad Nacional de Educación a Distancia), the Spanish national university of distance education. These two female students were following the course Inglés II para Turismo (English for Tourism Studies II), which is equivalent to a B2 level (CEFR 2001, 2007). They volunteered to participate in a project, which consisted in trying the app. Therefore, the data gathered derives from the pre- and post-questionnaire, as well as from their recordings. Also, one of the students exchanged two emails with the teaching team in which she showed her satisfaction and content with the task and stated that she had realized how much she liked AD. However, we will not take these emails as relevant. In this case, we consider that actual learning can be demonstrated by comparing the
results of the pre- and the post- questionnaire and also by comparing these to the students’ actual performance (their recording).

5.1. Sample and preliminary data: Results of the pre-questionnaire

The pre-questionnaire is meant to gather data on the users, such as linguistic background, previous knowledge of AD, etc., in order to assess the rate of improvement after using the app. In this case, it is remarkable that both voluntary students were bilingual: student A in Spanish and Galician, and student B in Spanish and Romanian. In both cases, they stated that they used English as an FL quite often, and student B pointed out that she used another third language: French. Regarding their previous knowledge of AD, student B had already heard of it, whereas student A had not. If students declare not to know what it is, they are invited to guess what they think it is. In this particular case, student A said that she thought it was a mobile app with videos to learn a language. Regarding the section where they had to translate words (from English to Spanish or vice versa), both of them provided correct answers, although student A left four out of ten terms without translation. Therefore, we can expect, out of these results, student B to be more motivated than student A to practice with VISP v1.

5.2. Results of the recording

Only one of the two students (student B) sent their recording, which is already a sign of (de)motivation. However, users are not obliged to send their recording, they can keep it for themselves and practice with AD as much as they want. Student A finished the task, as seen in the fact that she filled in the post-questionnaire, but decided not to send her recording. In what follows, the transcribed recording sent by student B is presented and analyzed:

(6)

Dark side of the Moulin Rouge at night, followed by a young black haired man, sitting in a dark room, in front of a typewriter, crying, with his hand in his forehead. The words on the paper come to life: “The woman I loved...” He makes a pause, and looks out of the window across the street, while the words start coming, and so he begins his story: “I first came to Paris, one year ago”.

Linguistically, the text in (7) is impeccable. The spoken discourse of the student was also perfect. The only errors we find are related to AD techniques themselves: she used too many words (76) for a clip of 31 seconds, due to the fact that she also included in the ADS the voice in off of the protagonist. Users are told, in the instructions, not to describe sounds or what is already heard, but still, this mistake is very common among students, as shown by Ibáñez Moreno and Vermeulen (2012,
As for the rest, we can say that the task performed by this student was very good, which is pointed out by the student herself in the post-questionnaire.

5.3. Results of the post-questionnaire

Both students filled in the post-questionnaire. Regarding answers to section 2.a, where they had to tick the expressions that were completely different from the ones they had used, they revealed that student A used all expressions differently, except of one (she ticked five boxes out of six). Student B ticked three boxes. These results are confirmed by the students’ answers to section 2.b (Any differences with what you said?, see illustration 3.2.b above), provided below:

(7)

Student A: “My sentences are shorter than these ones.”

Student B: “I think I described almost the same, just in a different way, with my own words, though I didn’t know it was in Paris (though I should’ve figured it out easily) and I also didn’t know his name was Christian”

Regarding student A, we cannot compare her statement to her actual performance because she did not send it, but regarding student B, who states that she described almost the same, we can see it is true. We can state, then, that student B shows high metacognitive skills. Also, her remark on the fact that she could not know about proper names is correct, and has been used for improving the post-questionnaire in future editions.

Both students stated that they did not use expressions such as we observe, we see... and that they used adjectives and adverbs that were different from the ones in the original ADS. Regarding students’ answers to section 3 (Reflections on AD, see illustration 5 above), student A answered 4 or 5, and student B answered 5 in all cases, so we can state that both students strongly believed that the tasks with AD were helpful for them to become aware of the others in communication, to try to find the accurate words and expressions, and to reflect upon their own language learning.

6. Conclusions and suggestions for future research

In this work we have presented the methodological steps taken by three members of the research group ATLAS to develop a MALL app that aims to help B1/B2 English language learners to work their oral skills, especially speaking. It has been developed departing from the idea that the use of AD, which offers the same information that is accessible visually in an oral way, can create an effective and motivating multimodal
learning environment. VISP v1, as a pilot version, was tested on two distance learning students of English. The results show the positive and promising potential of applying AD to a mobile app that aims at promoting oral skills in the FL, although there is still room for improvement.

We are already working on a second version, where we will implement more clips, for levels A1 and A2 tasks. We will also make it available in other languages, for instance for students of Spanish as an FL. In the near future we expect VISP to become a pioneering MALL app that uses AD as a technique to practice oral production skills in the FL. Another point of future research is the correlation between the user’s linguistic and cultural background and the successful implementation of VISP as a MALL app. By now, we can only say that student B—who spoke more languages, came from a multicultural environment (she was both Spanish and Romanian) and knew what AD is—was significantly more motivated from the beginning and showed her satisfaction with the app in several emails sent to the researchers. More data would contribute to shed some more light on this topic.

References


Bourne, Julian & Jiménez Hurtado, Catalina (2007). From the visual to the verbal in two languages: a contrastive analysis of the audio description of The Hours in English and Spanish, in Diaz Cintas J. et al. (eds): Media for All: Subtitling for the Deaf, Audio Description and Sign Language. Ámsterdam, Rodopi, 175-187.


Elorza, Izaskun, Castrillo, Maria Dolores, Ávila-Cabrera, José Javier, & Talaván Zanón, Noa (2014). Implementing situated m-learning activities for receptive oral skills. Presented at the International TISLID ’14 Conference, Ávila, May 7-9: http://cadiz.congresoseci.com/tislid14


Peskoe, Melany (2009). Descriptive Children’s Television: Bridging the Gap for Blind Kids While Benefiting All Kids, University of Louisville (on-line 03/03/2014).


Tight, Daniel (2010). Perceptual Learning Style Matching and L2 Vocabulary Acquisition, in Language Learning 60 (4), Language Learning Research Club, University of Michigan, 792-833.

