uc3m | Universidad Carlos III de Madrid

e-Archivo

This is the peer reviewed version of the following article:

García, B. (2017, 23 marzo). Bilingual subtitles for second-language acquisition and application to engineering education as learning pills. Computer Applications in Engineering Education, 25(3), 468-479,

which has been published in final form at https://doi.org/10.1002/cae.21814

This article may be used for non-commercial purposes in accordance with Wiley Terms and Conditions for Use of Self-Archived Versions.

© 2017 Wiley Periodicals, Inc.

Bilingual subtitles for second language acquisition and applied to engineering education as learning pills

Abstract

This paper introduces the concept of bilingual subtitles, a kind of captioning in which a pair of subtitles (in mother tongue, L1, and second language, L2) are shown at the same time on the screen. The aim of this dual subtitles is to help the final user in different learning processes, due to the fact that several capacities (listening, reading, and matching) are exercised at the same time by the learner while watching a dual captioned media. The contribution of this paper is three-folded. First, it presents DualSub, an open source desktop tool aimed to create bilingual subtitles. Second, a descriptive study has been designed and executed to evaluate the extent in which bilingual subtitles are perceived by final users in the incidental vocabulary knowledge of a second language. Third, an experimental case study in which dual subtitles are used in the engineering education arena has been also carried out. The results of these surveys confirm that bilingual subtitles are perceived as useful in the different dimensions of the incidental vocabulary learning process (form, meaning, use) and are also helpful applied to the educational domain (deliberate learning).

Keywords: Educational technology, subtitling, second-language acquisition, vocabulary knowledge, engineering education.

1 Introduction

A first language (also native language, mother tongue, arterial language, or L1) is the language a person has learned from birth. By contrast, a second language (or L2) is any language that one speaks other than one's first language [1]. The acquisition of a language is a natural process whereas learning a language is a conscious one [2]. The study of how a L2 is learned/acquired is referred to as second-language acquisition (SLA). Incidental learning occurs when learners acquire new aspects of their L2 without being focused on doing so. We can then distinguish incidental learning from deliberate learning. The latter takes place when we conduct an activity with the specific aim to learn something.

During the last decades, audiovisual materials (video, television programs, and so on) have been used as rich learning resources [3]. Although captions were originally developed for hearing-impaired persons in the 1980s and made accessible on television through the Teletext system, researchers quickly discovered that audiovisual material enhanced with captions or subtitles is a powerful pedagogical tool to improve the listening comprehension skills of learners [4][5][6].

Captioning and subtitling relate to the addition of onscreen text that renders dialogue. But both terms are significantly different. Subtitling is most frequently used as a way of translating a medium into another. Captioning, on the other hand, is more commonly used as a service to aid deaf and hard-of-hearing audiences. Usually, captions (also called closed captions) notate sound effects and other dramatically significant audio. Subtitling can be either interlingual, where the language of the television programs appears translated into the target language on the screen, or it can be intralingual, where it is usually targeted at a deaf audience, where the source language production is also used for the subtitles. The element of translation, which is present in interlingual and absent in intralingual distinguishes the two genres (Williams, 2000).

- L1 → L1 (intralingual)
- L2 → L2 (intralingual)
- L1 → L2 (interlingual)
- L2 → L1 (interlingual)

This piece of research introduces a novel topic on subtitling: the bilingual subtitles. In this kind of subtitles (also referred as dual subtitles), the language of the audiovisual material appears translated at the same time than the source language on the screen. The aim of bilingual subtitles is to comprise the benefits of intralingual and interlingual subtitles at once. On the one hand, intralingual subtitles provide the opportunity to have access to the written forms of the speech. Exposure to the written forms of words can be conducive to better performance on the vocabulary. On the other hand, interlingual subtitles requires listening attentively in order to recognize and fully absorb the content of the audiovisual material.

The contribution of this paper is three-folded. First, it presents DualSub, which is an open source tool specifically designed to produce bilingual subtitles. Second, a survey is performed to explore the value of dual subtitles for second-language learners. In this survey, the participants perform a self-assessment

judgment about various components of their own L2 incidental learning process with dual subtitles. Third, a case study to find out the usefulness of dual subtitles on deliberate learning has been carried out. This study has been focused on engineering education. Different groups of engineering students received an online master class in which different types of subtitles were used. Subsequently, several questions were asked of the students to find out the extend in which lesson had been understood.

The remainder of this paper is structured as follows. Section 2 reviews the relevant literature to provide the theoretical background for this study. Section 3 presents the research questions guiding this work study. Section 4 introduces DualSub, a desktop tool that allows to create bilingual subtitles for final users. Next, section 5 provides the design of the survey methodology including the instruments and participant selection. Data collection, and analysis of the studies are presented subsequently. The last sections conclude this work with discussion of findings, contributions, and suggestions for future research.

2 Literature review

2.1 Second language learning

Research on incidental learning has been mainly carried out in the area of reading [7][8], but also in the area of listening [9][10]. Learning a second language is a complex task than involves heterogeneous activities. According to the classic work by [11], a L2 learner has to face to four different problems:

- The problem of analysis. In order to learn a L2 the learner must analyze the stream of sounds and separate out its constituents.
- The problem of synthesis. In order to produce utterances that go beyond one-word sentences, the learner has to put the language constituents together.
- The embedding problem. Utterances are generally embedded in copious contextual information. Learning a language amounts to a gradual shift in the balance of linguistic and contextual information in favor of the former.
- The matching problem. Progress in language acquisition requires the learner to match continuously his own language performance against the standards of the target language speakers.

L2 learning largely depends on vocabulary, as the building blocks from which learners start their L2 acquisition [12]. During the past decades, L2 vocabulary learning has become of great research interest, especially when it comes to its incidental learning [13]. Gass [14] discusses key issues concerning incidental vocabulary learning, presenting an approach to incidental learning that draws attention to the recognition of syntactical categorization of the lexicon through context. The author suggests that vocabulary acquisition involve a certain degree of syntactic and lexical knowledge that learners heavily rely on for comprehension, and it is this relationship that needs to be taken into account for vocabulary learning purposes, whether it is incidental or intentional.

Research on vocabulary acquisition entails having an understanding of what "knowing a word" means so that one can use appropriate tools and procedures to measure vocabulary knowledge [15]. Vocabulary knowledge is a complex construct with multiple knowledge dimensions. One of the most influential descriptions of such aspects is shown as Table 1 [16]. This work provides a set of dimensions of what is involved in knowing a word regarding: form (spoken, written, and word parts), meaning (form and meaning, concept and referents, and associations), and use (grammatical functions, collocations, and constrains of use). L2 learners need to be aware of the fact that knowing a word involves both receptive knowledge (i.e. being able to retrieve a word's meaning through listening or reading its form) and also productive knowledge (i.e. being able to express meaning by using the accurate spoken or written word form).

Form	spoken	R What does the word sound like? P How is the word pronounced?
	written	R What does the word look like? P How is the word written and spelled?
	word parts	R What parts are recognizable in this word? P What word parts are needed to express the meaning?
Meaning	form and meaning	R What meaning does this word form signal? P What word form can be used to express this meaning?
	concepts and referents	R What is included in the concept? P What items can the concept refer to?
	associations	R What other words does this make us think of? P What other words could we use instead of this one?
Use	grammatical functions	R In what patterns does the word occur? P In what patterns must we use this word?
	collocations	R What words or types of words occur with this one? P What words or types of words must we use with this one?
	constraints on use	R Where, when, and how often would we expect to meet this word? P Where, when, and how often can we use this word??

R = receptive knowledge, P = productive knowledge

Table 1: Description of "what is involved in knowing a word" [16]

2.2 Subtitles as an aid for learning

We can find several methods of subtitled/captioned video for learning in the literature. The pioneer work to investigate the potential of captioning as a pedagogical tool was conducted by Price [4]. Approximately 500 participants watched four clips with or without captions. Findings showed that captioning resulted in superior comprehension independent of background, linguistic, and social variables. Since then, different studies have yielded encouraging results concerning the effectiveness of captioning for reinforcing L2 learners' listening comprehension [17][18] and vocabulary learning [19][20].

Verbatim subtitling refers to everything spoken including pause fillers, hesitations, etc. on the screen, whereas nonverbatim subtitling refers to summarized subtitles conveying only necessary information to get the film. Zarei [21] investigate the effect of verbatim and nonverbatim interlingual and intralingual subtitles on L2 vocabulary comprehension and production. This work revealed that nonverbatim subtitles resulted in more vocabulary comprehension regardless of whether they were interlingual or intralingual, whereas intralingual subtitles were more conductive to vocabulary production irrespective of whether they were verbatim or nonverbatim.

Perez presented a meta-analysis to synthesize primary research on the benefits of captioning on SLA. The overall results of this work support the claim that captioning helps learners to improve comprehension and fosters vocabulary acquisition [22].

As introduced before, vocabulary is the foundation of language learning. Researchers have found that knowledge of word meanings has a strong relationship to reading comprehension skill [23]. In a review of the literature published after the 1970s, the importance of vocabulary teaching has been reemphasized in language teaching [24]. These include translation and interpreting, the use of television and video for language learning and language teaching, the use of subtitles as an aid to second language acquisition and language learning and acquisition theory. The role of translation as a tool in language acquisition at an advanced level is discussed by Newmark [25]. The benefits of using television broadcasts and subtitles as an aid to SLA have been well documented [3][5][26].

2.3 Use of multimedia in education

According to Debevc et al. [27] video-based online lecture is one of the most powerful form for distance education since it can provide rich information, which is difficult to achieve through text, graphs, or verbally. From the students' point of view, video lectures provide a possibility to review and revise course materials at their better convenience. Moreover, video lectures give the opportunity to study for

those who are not able to participate in traditional classroom activities [28]. A comprehensive review of the use of media and communication technologies in e-learning activities can be found on [29].

In this domain, the use of Massive Open Online Courses (MOOCs) is supposing a revolution in education. Among other educational assets, this kind of courses are usually driven by video lessons. MOOCs has become in a hot topic in the educational research nowadays. Concerning engineering education, the number of MOOCs about mathematics, physics, statistics, and computer programming is increasing more and more [30]. For example, Muñoz-Merino et al. [31] show that students improved their grades significantly when using MOOCs technology when using an online educational platform with students in a remedial Physics course.

Moreover, the inclusion of multimedia contents is more and more used in teaching engineering courses, due to the fact that it is useful to help to deeply understand a problem (learning) and to apply knowledge to new problems to be solved (skills). This type of video tutorials, known by many authors as "learning pills" consists on small pieces of learning materials created as audiovisual content. In the broad sense, a learning pill can be seen as a simple exercise that summarizes some of the key concepts explained in class and promotes reflection and self-study [32].

The use of captioned video lectures constitutes a useful educational resource for students with disabilities. For example, Bengochea et al. discusses the use of subtitles in course videos as an accessible learning method to deaf students [33]. Nevertheless, the use of subtitles as a generic aid to engineering education has not fully addressed in the literature.

3 Objectives

From the literature presented in section 2 it is evident that research on the use of subtitling for learning purposes has been addressed. Nevertheless, research on bilingual subtitles is a novel topic, and so, this paper is aimed to fill this gap with several contributions.

First, a tool specifically designed to create dual subtitles has been designed, implemented and released as open source. Second, a survey has been carried out in order to assess the users' learning experiences with bilingual subtitles for incidental L2 learning. Third, another survey has been performed in order to assess if bilingual subtitles can help to engineering students to understand better a video learning pill in a second or foreign language (i.e. deliberate learning).

The broad questions driving the incidental L2 learning case can be stated as "Are these subtitles useful for L2 learners?" SLA is a complex process that involves many different aspects. As depicted in section 2, vocabulary is the foundation of language learning. In order to narrow the problem at hand, this research is focused in the learning process of L2 vocabulary. Thus, the first research question of this work is formulated as follows:

RQ1. Are dual subtitles useful to learn L2 vocabulary incidentally?

In order to find a valuable answer to this question, this work is based on the well-known definition of "what is involved in knowing a word" by Nation [16] (see Table 1 in section 2.1). Thus, three dimensions in a word knowing are used: form, meaning, and use. Therefore, RQ1 can be expressed as the following hypothesis:

H1. Dual subtitles are useful to learn the form of new vocabulary for L2 learners.
H2. Dual subtitles are useful to learn the meaning of new vocabulary for L2 learners.
H3. Dual subtitles are useful to learn the use of new vocabulary for L2 learners.

In dual subtitles, there are different configurations. First, the audio of the content can be in L1 and the subtitles in L1 and L2 (let's call this kind *intra-bilingual*). In the second configuration type the audio is L2 (let's call this kind *inter-bilingual*).

- L1 → L1+L2 (*intra-bilingual*)
- L2 → L1+L2 (*inter-bilingual*)

For this reason, the second research question related on SLA has to do with the preferred way of consuming dual subtitles:

RQ2. What is the preferred configuration for dual subtitles when learning a L2?

The second big objective of this paper is to study if bilingual subtitles can be a useful asset for deliberate learning, helping students to understand better a lesson. This part of the work is focused on engineering education, which can be seen as the activity of teaching theory and principles related to the practice of engineering profession.

As described in section 2, the use of multimedia in the classroom is becoming more and more popular. Common examples of these video lessons are the so-called "learning pills" and MOOCs. Moreover, the use of multimedia materials in foreign and second language (typically English) in engineering classrooms is becoming increasingly common. All in all, in addition to the inherent complexity of the subject under study, students need to make an extra effort to understand properly the content of this kind of video lessons in L2.

One simple possible solution to the problem would be using L1 applied to the video lesson. However, this solution applied to an engineering content can lead to potential misunderstanding of the lesson, due to lost-in-translation issues in the subject under study. In addition, the use of native L2 terms in non-L2 engineering classrooms is commonplace. Therefore, the joint use of these subtitles (L1+L2) in as bilingual subtitles for deliberate leaning of engineering subjects worth to be studied. This idea leads to the last research question of this work, which can be formulated as:

RQ3. Are dual subtitles useful applied to engineering education?

In order to narrow a bit this question, dual subtitles are applied to video lessons in form of learning pills for engineering students. Thus, RQ3 leads to the final hypothesis of this paper:

H4. Dual subtitles help to understand learning pills in engineering education.

4 DualSub, a tool for merging subtitles

Currently, there are plenty on media available with subtitles on different platforms (TV, Web, DVDs, etc). Nevertheless, the concept of dual subtitles is very scarce. In order to promote this kind of subtitles, a software application has been created: DualSub[1]. This tool has been released as open source (LGPL license), and it has been implemented with Java as a desktop application. For the shake of compatibility, the release distribution includes a wizard installer for Windows (.exe), a Mac OS X disk image (.dmg) installer, a Debian (.deb) package for Linux, and a cross-platform runnable Java archive (.jar). DualSub has been designed with the objective of creating dual subtitles easily for final users. To that aim, it has been implemented as a highly configurable tool from its GUI (see Figure 1).

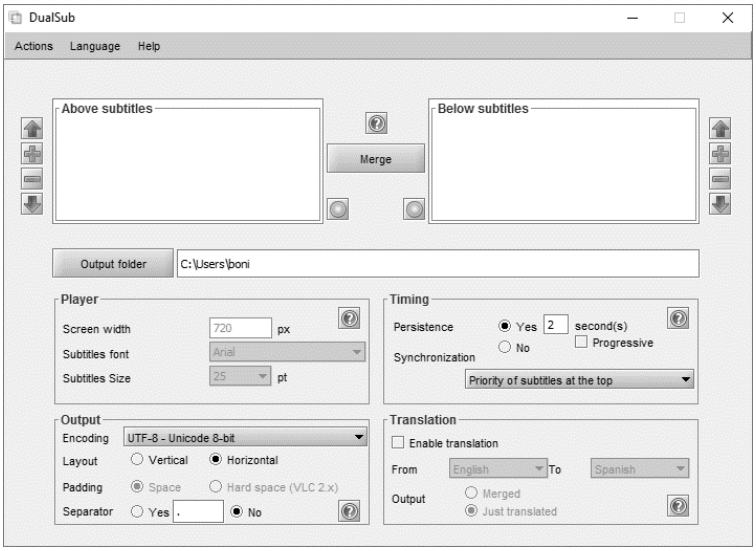

Figure 1: DualSub GUI

In order to create convenient dual subtitles, several aspects have to be taken into account. First of all, we need a pair of input subtitles, typically one in L1 and the other one in the L2. DualSub is a tool which manipulates theses subtitles. The format accepted for input subtitles in DualSub is SubRip (files with extension .srt), since it is the de-facto standard for the most of players nowadays. This kind of subtitles contain formatted lines of plain text in groups separated by a blank line. Subtitles are numbered sequentially starting at 1. The next line defines the time the subtitle should appear on the screen. After this timecode, the subtitle itself is shown in one or more lines. Finally, a blank line containing no text indicates the end of the subtitle. Given a pair of SubRip files, DualSub is able to merge them in a single subtitle file also in

[1] http://bonigarcia.github.io/dualsub/

SubRip format. The first challenge that DualSub faces is how to layout two subtitles at the same time. Let's consider the following example subtitles to be merged (in English and Spanish):

```
00:58:06,400 --> 00:58:13,400
and our desire to understand and
explore this beautiful universe.
```

```
00:58:06,400 --> 00:58:13,400
y nuestro deseo de comprender y
explorar este hermoso universo.
```

There are two main possibilities to show these two subtitles at once: using horizontal and vertical layout. DualSub support both layouts out of the box. On the one hand, in the horizontal layout the available screen width is divided in two columns in each subtitle is shown (Figure 2-a). In order to place correctly the subtitles in each column, DualSub is able to measure the width occupied by the subtitles lines, using spaces as padding. Nevertheless, some players trim the space characters at the beginning and the end of SubRip subtitles. For that reason, DualSub can be configured to use a separator character (a dot by default) to avoid these potential problems (Figure 2-b). Regarding the vertical layout, both subtitles are placed one over the other (Figure 2-c), with the possibility of adding some separator between the two subtitles (Figure 2-d).

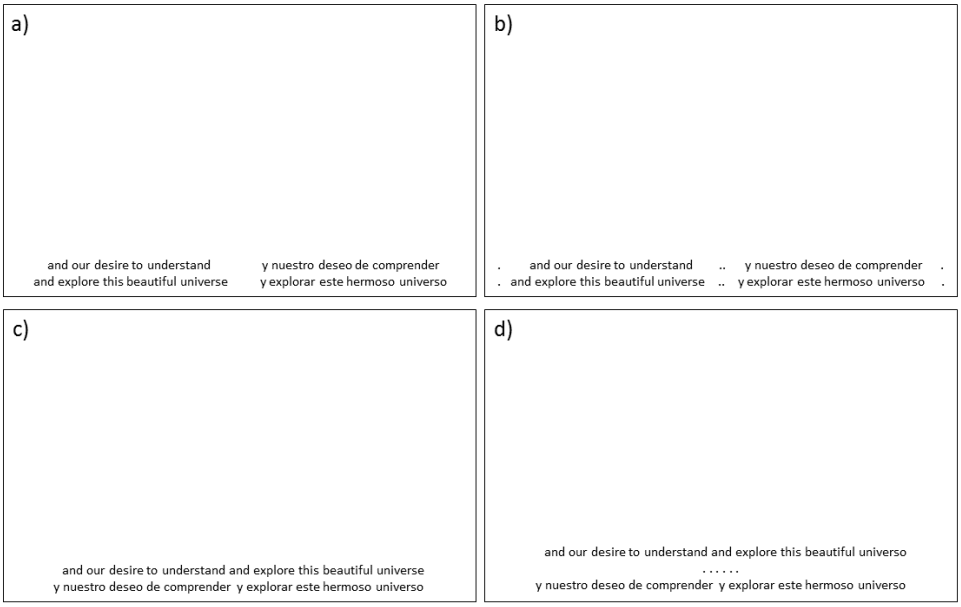

Figure 2: Dual subtitles layouts

In this example, the timing of both subtitles (English and Spanish) starts and ends at the same time. This is not a requirement to merge the subtitles correctly, since DualSub is able to recognize the time slots occupied by subtitles, creating the resulting dual subtitles conveniently. In order to avoid overlapping problems (i.e. that too many subtitles grouped together), DualSub offers the possibility of prioritizing the time selection in the resulting dual subtitles (i.e. priority of L1, L2, maximize, or minimize the overlapping time).

Dual subtitles show the double of information on the screen at the same time. Therefore, it is likely that users are not able to read all the text in the subtitle time. In order to help users to read the most of the dual subtitles, DualSub implements a useful feature for bilingual subtitles named *persistence*. This feature allows to hold the subtitles a configurable number of seconds even when the theoretical subtitle time is over. In addition, this extra time can be progressive, i.e., wait a configurable number of seconds for each line of input subtitle.

Another out of the box feature of DualSub is the support for multiple charsets. Due to the fact that different languages uses different encodings, DualSub is able to automatically find out the charset of the input subtitles. Then, the output encoding should be explicitly configured by user. DualSub supports up to 158 different charsets (e.g. UTF-8, UTF-16, UTF-32, ISO-8859-1, among others).

Last by not least, DualSub allows to translate subtitles. Sometimes, it can be very difficult to find subtitles in two languages for the same content. For this reason, DualSub allows to translate a single input subtitle, merging the translation with the original, if desired. To accomplish this feature, DualSub relies on the online service Google Translate. Although this service is not 100% accurate, it is a very good alternative when users only have a single input subtitle.

5 Methodology

5.1 Instruments

Given the research questions stated in section before, two different surveys have been carried out. The objectives of each surveys are the following:

1. To assess the incidental vocabulary for L2 learners
2. To assess the deliberate learning of engineering subjects when consuming online lessons with dual subtitles

Regarding the first survey, and as depicted in section 2.1, each dimension of the vocabulary acquisition by Nation (form, meaning, and use) are divided in two categories (respectively: spoken, written, word parts, form and meaning, concepts and referents, associations, grammatical functions, collocations, and constraints on use). And each category can be related to the productive (P) or receptive (R) knowledge. All in all, there are 18 questions, and for this reason the instrument designed in this survey to address RQ1 needs to provide an answer for each question. Thus, the instrument of this survey is a self-assessment questionnaire which contains 18 items, one by each category and type of the vocabulary dimensions of Nation. To elicit the required data, a 5-point Likert-type scale questionnaire was used, ranging from "strongly disagree" through "neither agree or disagree" to "strongly agree".

To provide the logical dimensions of validity before its statistical dimension, as Black and Champion [34] recommended, the first draft of the instrument was revised based on opinions of experts, two language teachers and two educational science professors. Moreover, a pilot study was carried out before launching the final questionnaire. Pilot studies are intended to identify any problems with the questionnaire itself, as well as with the response rate and follow-up procedures [35]. A group of 12 DualSub users was selected to test the new version of DualSub and to carry out the evaluation of the questionnaire. This group is made up for those users who probe some interest in dual subtitle by reporting any issue in the DualSub forum[2], or contacted the author to request some feature or comment. As a result, some items were revised according to the results from the pretest and pilot tests, to improve face and content validity, as well as reliability. The final version of the questionnaire is in Appendix.

In order to find an answer to RQ2, several questions are included in the final version of the questionnaire, namely:

1. What is your motivation to use dual subtitles?
2. In which language do you listen your movies/series with dual subtitles?
3. When you watch movies/series with dual subtitles, what subtitles do you prefer to read?
4. When you listen/read a word you do not understand, what subtitles do you prefer to read?
5. What is your level in the second language?

The possible answers for the first of these questions are the following:

- To learn a second language
- To watch the same contents with people who speak a different language
- Accessibility reasons (deafness, etc)
- Other reasons

Regarding questions 2 to 4, participants can select between mother tongue (L1), the second language to be learnt (L2), or other (open answer). Regarding question 5, the possible answers are beginner (A1), elementary (A2), intermediate (B1), upper intermediate (B2), advanced (C1), and proficiency (C2).

In order to address RQ3, a second survey focused on engineering education has been designed and executed. The objective of this survey was to analyze the extend in with the students understand a video lesson in form of learning pill. This audio of lesson was L2, while the students' language is different (i.e. L1). The process to analyze the value of dual subtitles is the following

1. Students are asked to watch the video lesson (i.e. the learning pill) in L2 without any subtitle. Immediately, students should make a small test (i.e. a group of questions) about the video lesson.
2. A week later, the same group of students are asked to watch the same video lesson, but this time the video lesson is captioned with dual subtitles (L1+L2). Just like in the step before, the same test should be completed by the students after viewing the lesson.

[2] https://groups.google.com/forum/?hl=es#!forum/dualsub

The objective here is to compare the results from step 1 and step 2. The differences between these two steps is that in the second experiment, dual subtitles are present in the lesson. If the use of this dual subtitles have a significant impact on the test results, the null hypothesis for H4 (i.e., dual subtitles *do not* help to understand learning pills in engineering education) could be rejected.

5.2 Participants

The first survey was conducted using the version 1.1.0 of DualSub. The first version of DualSub (i.e. version 1.0.0) had over than 10K downloads[3] world-wide. With this critical mass of users, the next version includes the questionnaire designed to address the proposed research questions. DualSub 1.1.0 GUI asks for user participation in the survey, showing a link to the online poll hosted on Google Forms.

It is not enough to decide how many people to survey: the target population and the valid sample should be defined. On the one hand, the target population for a survey is the group or the individuals to whom it applies. On the other hand, a valid sample is a representative subset of the target population [36]. The target population for this survey is the users of DualSub 1.1.0. The specific target population is defined by means of an exclusion criterion, which is determined with the answer to the question of the poll: "*What is your motivation to use dual subtitles?*". Only the people whom answer is the first option (i.e. to learn a second language) are considered as the valid sample to the survey analysis.

The second survey has been carried out in the engineering degree at University Center for Technology and Digital Art (U-tad) in Spain. The second-grade students of the subject "Computer networks" were required to take a learning pill about networking. The subject of this lesson was "UDP hole punching", which is a commonly used technique employed to establish a peer to peer communication behind network address translator (NAT) devices [37]. This topic was introduced in the classroom, but the low-level details of this mechanism were unknow to the students. As depicted in section 5.1, the students are required to take the learning pill about this topic without and with bilingual subtitles. This learning pill is a short video lesson of 5 minutes available on YouTube. The language of the lesson speaker is English. Therefore, the representative sample for this survey is those students whom mother tongue is not English. The test to be done for students is a group of 16 questions about the content of the video lesson.

6 Data analysis

6.1 Vocabulary learning

The data of this survey has been analysed using the Statistical Package for the Social Sciences (SPSS) version 20. First of all, in order to check the validity of the questionnaire, the Kaiser-Meyer-Olkin (KMO) and Bartlett's Test of Sphericity (BTS) were analysed. As shown in Table 2, the KMO value was found to be 0.875 which can be said to be a great value and BTS showed an approximately Chi-Square value of 1502.494 ($p < 0.000$), which meant that the correlation matrix of data is appropriate [38].

Kaiser-Meyer-Olkin measure of sampling adequacy	.875
Barlett's Test of Sphericity	
Approximate χ^2	1502.494
df	153
Sig.	.000

Table 2: KMO and BTS results

Reliability was examined using Cronbach's alpha values for each variable. The alpha statistic measures the internal consistency reliability among a group of items that combine to form a single scale. It indicates how well the different items complement each other in their measurement of different aspects of the same variable or quality [35]. Cronbach's alpha coefficient regarding the 18 items constituting the instrument was calculated as 0.997. Cronbach's alpha coefficient values higher than .70 are considered as good, and when the value is close to 1.00, it is considered excellent [39]. Therefore, we conclude that the collected data are excellent in terms of reliability.

At the time of writing of this paper DualSub 1.1.0 has been downloaded 4238 times. Each user is a potential participant of the survey, but only 94 questionnaire responses were collected (a response rate of

[3] https://api.github.com/repos/bonigarcia/dualsub/releases

2.21%). Filtering the participants by means of the motivation for using DualSub as defined in section 5.2 (i.e., using DualSub to learn a L2), the resulting valid responses was 62.

The mean age for these 62 participants was 32.5 years with a standard deviation of 13.6 years. Regarding gender, the rate was 95.161% men and 4.838% women, from 29 different countries (see Figure 3). There are 20 different mother tongue of the participants, being Arabic and Spanish the first and second position respectively, and the predominant L2 is English (see Figure 4).

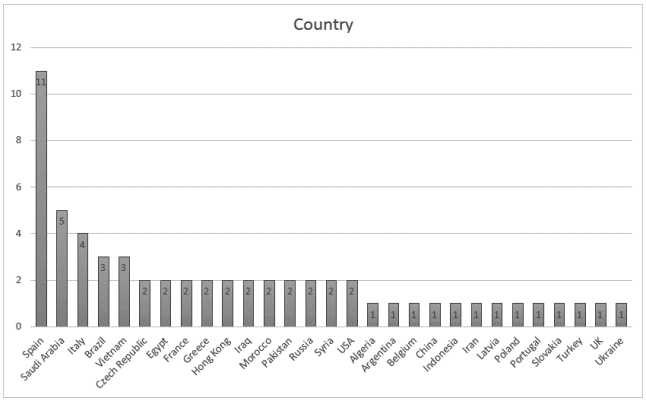

Figure 3: Survey participant's origin countries

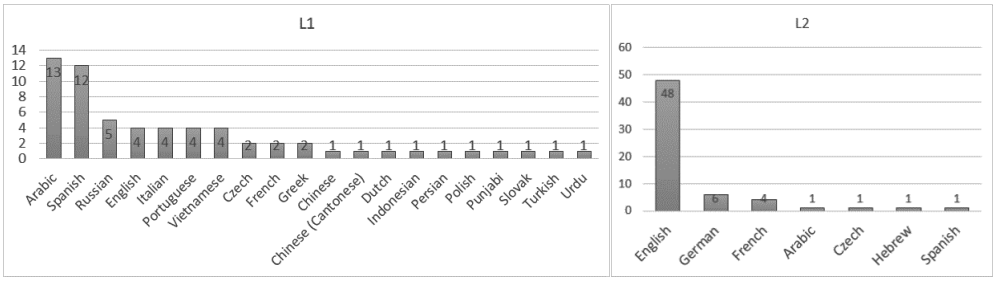

Figure 4: Survey participants' L1 and L2

Table 3 presents a summary of the descriptive statistics (minimum, maximum, mean, and standard deviation) for the each of the 18 items (see Appendix) of the survey. In the next section an interpretation of these figures is provided.

Item id	Minimum	Maximum	Mean	Std. Deviation
F1-R	2	5	4.06	1.022
F1-P	1	5	3.77	1.151
F1-P	1	5	3.77	1.151
F2-R	1	5	3.84	1.176
F2-P	1	5	3.84	.995
F3-R	2	5	3.92	.980
F3-P	2	5	3.85	.938
M1-R	2	5	3.98	.949
M1-P	2	5	3.85	.884
M2-R	1	5	3.85	.921
M2-P	2	5	3.82	.915
M3-R	1	5	3.79	.977
M3-P	1	5	3.65	.907
U1-R	2	5	3.87	.877
U1-P	1	5	3.95	.999
U2-R	1	5	3.94	.973
U2-P	1	5	3.95	1.015
U3-R	1	5	3.77	1.108
U3-P	1	5	3.90	1.020

Table 3: Survey items descriptive statistics

Finally, the Pearson correlation of the questionnaire items is shown in Table 4. These data show a positive correlation among all variables, which means that every construct moves in the same direction. Moreover, there are several correlation coefficients above 0.80, showing a strong correlation between variables these [40]. This means that there is strong linear relationship among some variables, especially from the vocabulary dimensions meaning and use.

	F1-R	F1-P	F2-R	F2-P	F3-R	F3-P	M1-R	M1-P	M2-R	M2-P	M3-R	M3-P	U1-R	U1-P	U2-R	U2-P	U3-R	U3-P
F1-R																		
F1-P	.626																	
F2-R	.636	.663																
F2-P	.526	.554	.776															
F3-R	.594	.652	.586	.625														
F3-P	.659	.546	.706	.659	.682													
M1-R	.660	.612	.732	.657	.774	.752												
M1-P	.627	.676	.655	.718	.800	.784	.798											
M2-R	.620	.711	.690	.672	.732	.678	.748	.920										
M2-P	.573	.615	.643	.598	.770	.753	.695	.839	.825									
M3-R	.588	.715	.655	.622	.786	.663	.756	.913	.931	.783								
M3-P	.432	.534	.560	.517	.594	.593	.621	.630	.664	.654	.654							
U1-R	.594	.587	.678	.595	.750	.714	.883	.821	.829	.788	.809	.807						
U1-P	.629	.646	.621	.553	.632	.605	.673	.716	.848	.708	.762	.686	.797					
U2-R	.614	.719	.678	.633	.717	.636	.744	.751	.813	.687	.779	.735	.854	.823				
U2-P	.603	.748	.707	.576	.787	.595	.816	.777	.817	.697	.816	.675	.821	.741	.877			
U3-R	.549	.641	.626	.532	.738	.678	.636	.803	.852	.833	.819	.718	.779	.746	.823	.792		
U3-P	.619	.679	.684	.518	.714	.636	.760	.784	.841	.737	.835	.724	.865	.832	.886	.882	.880	

Table 4: Pearson correlation results

6.2 Engineering education

This section describes the results of the experiment on engineering education using learning pills with dual subtitles. This test was optative for students, but those students completing the survey were granted with a 0.3 points in the final qualification. The number of collected answers were 19 out of 29 total students (response rate of 65.51%). All of these 19 students were not native English speakers (all of them were Spanish).

As described on section 3, this survey has two parts. First, the students watch the video lesson in L2 without subtitles and carry out a test about the content of the learning pill. Second, the same students watch the same video lesson in L2 but this time with dual subtitles, repeating the same test again. The professor qualified both tests for each student. The descriptive statics for the qualifications obtained by the students in both stages are depicted in the following table:

Experiment	Mean	Std. Deviation
First stage: learning pill in L2 without subtitles	6.7105	.43908
Second stage: learning pill in L2 with dual subtitles	8.1140	.54014

Table 5: Learning pill descriptive statistics

To analyze the statistical significance of this result, a paired t-test was used, yielding a p-value of 0.007. This value is lower than the threshold 0.01, and therefore the null hypothesis for H4 can be rejected. This fact confirms that there is a positive impact on the performance of students when dual subtitles are available in the video lesson.

7 Discussion

As depicted in section 3, regarding the first research question (RQ1: Are dual subtitles useful to learn L2 vocabulary?) three hypotheses were stated (H1, H2, H3). These hypotheses are linked to what it is

involved in knowing a word defined by Nation, 2001. In other words, the dimension form, meaning, and use are tested, both in the productive and the receptive body of knowledge.

Figure 5 presents a radar diagram in which each axis in one the questionnaire items. The values are taken from the descriptive analysis presented in Table 3. Two different areas are shown in this diagram: the productive and receptive knowledge perceived by DualSub users when using dual subtitles. The Likert scale used in this survey varies from 1 ("strongly disagree") to 5 ("strongly agree"). Therefore, the area of chart in Figure 5 describes the level of usefulness of dual subtitles perceived by DualSub users. The greater this area is, the better. The chart shows that the perception of dual subtitles by users is always around 4 (i.e. "agree" in terms of the Likert scale) for each of the items, both productive and receptive.

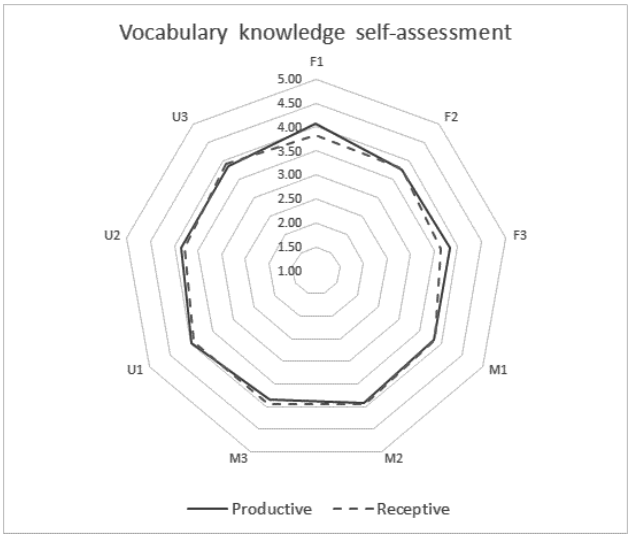

Figure 5: Vocabulary knowledge self-assessment result

In the light of this result, we can conclude that the hypothesis H1, H2, and H3 (RQ1) are meet, since the survey shows that the perception of DualSub users are positive for each category in the dimensions involved in learning new words by watching media with dual subtitles.

Regarding the second research question (RQ2: What is the preferred configuration for dual subtitles when learning a L2?), as introduced in section 1, various combinations of language transfer may be practiced during subtitling, depending on the requirements of the particular students (L1 → L1+L2 or L2 → L1+L2). For this reason, the questionnaire includes questions about the way of consuming the dual subtitles. The results of these questions summarized in Table 6. On the one hand, these data show that the preferred subtitles read by DualSub users are the subtitles in L2 (82.26%). On the other hand, the major part of DualSub users (77.42%) in this survey listen to L2 for media with dual subtitles. Finally, a significant rate of users read L1 when a new word appears in the media (69.35%). Therefore, we can conclude that the preferred option to consume dual subtitles is L2 → L1+L2 (inter-bilingual), reading the L2 subtitles while watching the media and switching to L1 when for new words.

	L1	**L2**
In which language do you listen your movies/series with dual subtitles?	17.74%	82.26%
When you watch movies/series with dual subtitles, what subtitles do you prefer to read?	22.58%	77.42%
When you listen/read a word you do not understand, what subtitles do you prefer to read?	69.35%	30.65%

Table 6: Dual subtitles use results

Another conclusion we can draw has to with the L2 level of the users that consumes dual subtitles. In the light of the results (see Figure 6), dual subtitles are used mainly for first stage learners (beginner to intermediate), less by upper intermediate and advance users, and it is not needed (as expected) by proficiency users.

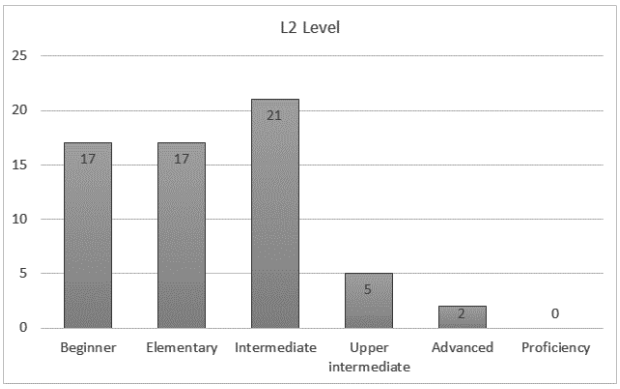

Figure 6: Action taken because of a new word by DualSub users

Regarding RQ3, the results presented in section 6.2 confirms that the null hypothesis for H4 (i.e., dual subtitles *do not* help to understand learning pills in engineering education) can be rejected, and therefore the alternative hypothesis H4 (dual subtitles help to understand learning pills in engineering education) can be accepted.

8 Conclusions

This piece of research has introduced the concept of bilingual (or dual) subtitles. This kind of subtitles encompasses all of the benefits of using video, subtitles, listening, and translation at the same time. The combination of aural, visual and written elements required to subtitle competently makes it unique for learners in L2.

This work has analysed the impact of dual subtitles in different fields. First, a survey has been carried out to find out if dual subtitles help in the L2 acquisition incidental process. In this domain, vocabulary forms the biggest part of the meaning of any language, and it is the biggest problem for most learners. Bilingual subtitles contribute to vocabulary production significantly better than intralingual and/or interlingual subtitles, due to the fact that having the translation available, users are aware of the grammatically correct, well-punctuated and unambiguous written form in both L1 and L2.

An open source tool called DualSub has been developed to help users to create dual subtitles. A questionnaire has been designed and distributed to the DualSub users. This instrument has been designed to evaluate whether or not the final users perceives dual subtitles as useful in the acquisition of new L2 vocabulary. The self-assessment results confirm that users find useful this kind of subtitles in the different dimension of the vocabulary incidental learning (form, meaning, and use). Moreover, this study has revealed that the favorite way of consuming dual subtitles is by means of inter-bilingual subtitles (L2 → L1+L2, i.e., audio in L2 with dual subtitles), reading L2 by default, and switching to L1 when new words appears on the screen.

Finally, a second case study of dual subtitles has been carried out. This study is focused on the engineering education, in which the use of dual subtitles can be used to help students to understand video lessons in L2. In the light of the results, dual subtitles have been revealed as a useful tool in learning pills in engineering, concretely in a networking classroom. The improvement in the qualifications of the video lesson was statistically significant when dual subtitles were used.

The future work for this research line can be continued by means of more experimental research. For example, assessing the degree of L2 vocabulary acquired by groups of users watching different configurations of dual subtitles (inter or intra-bilingual). Moreover, different surveys can be carried out the deliberate learning domain, for example integrating video lesson with dual subtitles in different kinds of MOOCs.

Acknowledges

I would like to acknowledge the participants of the surveys that made possible this paper. First, thanks to those DualSub users who completed the survey about incidental L2 vocabulary learning. Second, thanks to the second-year students of Computer Networks of the academic year 2016/2017 at University Center for Technology and Digital Art (U-tad) in Spain for participating in the case study on dual subtitles applied to engineering education.

Appendix. Vocabulary learning questionnaire items

Bilingual subtitles are useful for second-language acquisition because I learn…
Form
F1-R: to recognize a word when I hear
F1-P: to say a word with correct pronunciation including stress
F2-R: to recognize the written form of a word when it is met in reading
F2-P: to write a word with correct spelling
F3-R: to recognize a word that it is made up of the parts and being able to relate these parts to its meaning
F3-P: to construct a word using the right word parts in their appropriate forms
Meaning form and meaning
M1-R: to know that a word signals a particular meaning
M1-P: to produce a word in different contexts to express the range of meanings of the word
M2-R: to know the concept behind a word which will allow understanding in a variety of contexts
M2-P: to produce a word in different contexts to express the range of meanings
M3-R: to know that there are related words
M3-P: to produce synonyms and opposites
Use
U1-R: to know what a word means in the particular context in which it has just occurred
U1-P: to use a word correctly in an original sentence
U2-R: to recognize that a word has been used correctly in the sentence in which occurs
U2-P: to produce words that commonly occur with it
U3-R: to know that a word is not an uncommon word and is not a pejorative word
U3-P: to decide to use or not use a word to suit the degree of formality of the situation

References

[1] L. Bloomfield, "Language, motilal banarasidass publication pvt," Ltd, 1994.

[2] S. Krashen, *Principles and practice in second language acquisition*, (1982).

[3] R. Vanderplank, *Déjà vu? A decade of research on language laboratories, television and video in language learning*, Language teaching 43 (2010), no. 01, 1-37.

[4] K. Price, *Closed-captioned tv: An untapped resource*, Matsol Newsletter 12 (1983), no. 2, 1-8.

[5] M. Danan, *Captioning and subtitling: Undervalued language learning strategies*, Meta: Journal des traducteursMeta:/Translators' Journal 49 (2004), no. 1, 67-77.

[6] H. Williams and D. Thorne, *The value of teletext subtitling as a medium for language learning*, System 28 (2000), no. 2, 217-228.

[7] B. Dupuy and S. D. Krashen, *Incidental vocabulary acquisition in french as a foreign language*, Applied Language Learning 4 (1993), 55-63.

[8] J. H. Hulstijn, "Retention of inferred and given word meanings: Experiments in incidental vocabulary learning," *Vocabulary and applied linguistics*, Springer, 1992, pp. 113-125.

[9] H. Van Zeeland and N. Schmitt, *Incidental vocabulary acquisition through l2 listening: A dimensions approach*, System 41 (2013), no. 3, 609-624.

[10] K. Vidal, *Academic listening: A source of vocabulary acquisition?*, Applied linguistics 24 (2003), no. 1, 56-89.

[11] W. Klein, *Second language acquisition*, Cambridge University Press, 1986.

[12] R. Carter and M. McCarthy, *Vocabulary and language teaching*, Routledge, 2014.

[13] R. Ramos and F. Dario, *Incidental vocabulary learning in second language acquisition: A literature review*, Profile Issues in TeachersProfessional Development 17 (2015), no. 1, 157-166.

[14] S. Gass, *Discussion*, Studies in second language acquisition 21 (1999), no. 02, 319-333.

[15] P. Bogaards, *Testing l2 vocabulary knowledge at a high level: The case of the euralex french tests*, Applied Linguistics 21 (2000), no. 4, 490-516.

[16] I. S. Nation, *Learning vocabulary in another language*, Ernst Klett Sprachen, 2001.

[17] T. J. Garza, *Evaluating the use of captioned video materials in advanced foreign language learning*, Foreign Language Annals 24 (1991), no. 3, 239-258.
[18] N. Guichon and S. McLornan, *The effects of multimodality on l2 learners: Implications for call resource design*, System 36 (2008), no. 1, 85-93.
[19] M. Danan, *Reversed subtitling and dual coding theory: New directions for foreign language instruction*, Language learning 42 (1992), no. 4, 497-527.
[20] T. Sydorenko, *Modality of input and vocabulary acquisition*, Language Learning & Technology 14 (2010), no. 2, 50-73.
[21] A. A. Zarei and Z. Rashvand, *The effect of interlingual and intralingual, verbatim and nonverbatim subtitles on l2 vocabulary comprehension and production*, Journal of language teaching and research 2 (2011), no. 3, 618-625.
[22] M. M. Perez, W. Van Den Noortgate and P. Desmet, *Captioned video for l2 listening and vocabulary learning: A meta-analysis*, System 41 (2013), no. 3, 720-739.
[23] S. A. Stahl and M. M. Fairbanks, *The effects of vocabulary instruction: A model-based meta-analysis*, Review of educational research 56 (1986), no. 1, 72-110.
[24] M. McCarthy, *Recent directions in vocabulary teaching*, The Language Teacher 14 (1990), no. 12, 9-11.
[25] P. Newmark, *About translation*, vol. 74, Multilingual matters, 1991.
[26] S. B. Neuman and P. Koskinen, *Captioned television as comprehensible input: Effects of incidental word learning from context for language minority students*, Reading research quarterly (1992), 95-106.
[27] M. Debevc, R. Šafarič and M. Golob, *Hypervideo application on an experimental control system as an approach to education*, Computer Applications in Engineering Education 16 (2008), no. 1, 31-44.
[28] A. Kybartaite, J. Nousiainen and J. Malmivuo, *Technologies and methods in virtual campus for improving learning process*, Computer Applications in Engineering Education 21 (2013), no. 1, 185-192.
[29] D. Gañán, S. Caballé, J. Conesa, L. Barolli, E. Kulla and E. Spaho, *A systematic review of multimedia resources to support teaching and learning in virtual environments*, Complex, Intelligent and Software Intensive Systems (CISIS), 2014 Eighth International Conference on, IEEE, 2014, pp. 249-256.
[30] S. Iqbal, X. Zang, Y. Zhu, Y. Y. Chen and J. Zhao, *On the impact of moocs on engineering education*, MOOC, Innovation and Technology in Education (MITE), 2014 IEEE International Conference on, IEEE, 2014, pp. 101-104.
[31] P. J. Muñoz-Merino, J. A. Ruipérez-Valiente, C. Delgado Kloos, M. A. Auger, S. Briz, V. de Castro and S. N. Santalla, *Flipping the classroom to improve learning with moocs technology*, Computer Applications in Engineering Education (2016).
[32] M. Muñoz-Organero, P. J. Muñoz-Merino and C. D. Kloos, *Sending learning pills to mobile devices in class to enhance student performance and motivation in network services configuration courses*, IEEE transactions on Education 55 (2012), no. 1, 83-87.
[33] L. Bengochea and F. Budia, *Subtitled video tutorials, an accessible teaching material*, Journal of accessibility and design for all 2 (2012), no. 2, 155-164.
[34] J J. A. Black and D. J. Champion, *Methods and issues in social research*, John Wiley & Sons, 1976.
[35] B. Kitchenham and S. L. Pfleeger, *Principles of survey research part 4: Questionnaire evaluation*, SIGSOFT Softw. Eng. Notes 27 (2002), no. 3, 20-23.
[36] B. Kitchenham and S. L. Pfleeger, *Principles of survey research: Part 5: Populations and samples*, ACM SIGSOFT Software Engineering Notes 27 (2002), no. 5, 17-20.
[37] P. Srisuresh, B. Ford and D. Kegel, "State of peer-to-peer (p2p) communication across network address translators (nats)," 2008.
[38] G. D. Hutcheson and N. Sofroniou, *The multivariate social scientist: Introductory statistics using generalized linear models*, Sage, 1999.
[39] J. M. Cortina, *What is coefficient alpha? An examination of theory and applications*, Journal of applied psychology 78 (1993), no. 1, 98.
[40] F. Hair, R. E. Anderson, R. L. Tatham and W. C. Black, *Multivariate data analysis . Englewood cliff*, New Jersey, USA (1998).