The Riches of Hands-on Subtitling in the Foreign Language Classroom

Jorge Díaz Cintas[1] · 🆔 https://orcid.org/0000-0002-1058-5757

University College London

Gower Street, London, WC1E 6BT, UK

Chengcheng Wang · 🆔 https://orcid.org/0000-0003-1948-9363

University College London

Gower Street, London, WC1E 6BT, UK

ABSTRACT

Using subtitles and subtitling as a means of diversifying foreign language teaching and learning has become increasingly popular in recent decades, particularly across Europe, where the European Commission has promoted, among others, the development of projects like ClipFlair, a web-based subtitling platform for foreign language learning (FLL). As part of this surge in research on FLL through subtitling, this empirical study was conducted in mainland China, where the role of subtitling in foreign language classroom has not been widely recognised by scholars. The experiment, carried out with seventeen higher education level Chinese L1 students, explored the effects of performing subtitling activities on English L2 vocabulary acquisition and discovered that doing subtitling tasks from L2 to L1 can result in a significantly better performance in vocabulary acquisition than doing intralingual subtitling activities (L2 to L2) or doing non-subtitling activities.

Keywords: subtitling, translation, foreign language learning, vocabulary acquisition, second language acquisition.

RESUMEN

El uso de los subtítulos y de la subtitulación como prácticas didácticas que ayudan a diversificar la enseñanza y aprendizaje de lenguas extranjeras ha crecido en popularidad en las últimas décadas, especialmente en Europa, donde la Comisión Europea ha fomentado, entre otros, proyectos como ClipFlair, una plataforma de subtitulado en la nube para el aprendizaje de lenguas extranjeras. Enmarcado dentro de este auge investigador, el estudio empírico que aquí se documenta se ha llevado a cabo en China, donde el uso de la subtitulación en el aula de idiomas extranjeros apenas si ha recibido atención por parte de los académicos. El experimento, que ha contado con diecisiete estudiantes de grado y chino como L1, explora el impacto de la subtitulación en la adquisición de vocabulario en inglés (L2). Los resultados demuestran que la práctica activa de subtitular videos interlingüísticamente, de L2 a L1, puede conducir a una mayor adquisición de vocabulario que cuando las tareas se centran en la subtitulación intralingüística (de L2 a L2) o son actividades que no tienen nada que ver con la subtitulación.

Palabras clave: subtitulado, subtitulación, traducción, enseñanza de lenguas extranjeras, adquisición de vocabulario.

1. Introduction

For the past few decades, the role of subtitles as a means of promoting foreign language learning (FLL) has been explored enthusiastically by scholars from all over the world. Europe, in particular, has become a vigorous centre for the study of subtitles in the foreign language (FL) classroom, thanks, among others, to the active promotion spearheaded by the European Commission "to develop multilingualism" (Gambier, 2015: 64). Early evidence by scholars such as Vanderplank (1988, 1990), Danan (1992) and d'Ydewalle and Van de Poel (1999) has shown

---

that subtitled videos, by combining sounds, images, social-cultural information and text, can "facilitate formal and incidental language learning" (Danan, 2015: 41).

The use of audiovisual materials can be considered an important component in computer-assisted language learning (Brett and González-Lloret, 2009). One of the advantages of using subtitles in FLL derives from the so-called dual coding theory (Paivio, 1986). By providing visual and verbal codes concurrently, subtitled productions trigger multiple cognitive connexions in the viewers/readers' minds while processing the information, which is then stored in their memory and facilitates a reinforced retrieval at a later stage. That is, relying on both visual and verbal codes can boost performance in subsequent information recall. In addition, subtitles enable students to "see and appreciate how native speakers interact in everyday conversations" (Díaz Cintas and Fernández Cruz, 2008: 202).

Learning a FL through watching subtitled videos is not devoid of challenges since, as highlighted by Neves (2004: 129), students are often "taken on roller coaster rides" along images, sounds and texts that "come and go at a pace that cannot be altered". Learners are forced to sharpen their senses, if they are to keep in track with the multitude of signs to be decoded, especially considering that the sounds and/or texts "come in a code that is not fully mastered: a foreign language" (ibid.). Therefore, it is not surprising that in some experiments "educational benefits came only to those who put in considerable mental efforts" (Vanderplank, 2015: 23).

The traditional approach to exploiting subtitled materials, i.e., students watching the videos with subtitles and listening to the original soundtrack, has been deemed to be too passive an activity and, in recent years, researchers have advocated a more dynamic exploitation of subtitles by endorsing activities that require students to produce their own subtitles (Talaván, 2006, 2010; Incalcaterra Mcloughlin and Lertola, 2011; Sokoli *et al*., 2011; Lertola, 2013; He, 2015). Adopting the same philosophy, the research reported in this paper is the result of an investigation into the effect on vocabulary acquisition of producing subtitles. To this end, an experiment was conducted with seventeen students in higher education, divided into three groups doing (1) intralingual subtitling activities (English to English), (2) interlingual subtitling activities (English to Chinese), and (3) traditional reading and writing activities. Test results and questionnaires were collected for data analysis, with the ultimate goal of gauging and comparing their performance in vocabulary acquisition through these various activities.

2. Subtitling in the foreign language classroom

The educational potential of subtitles was not tapped into by researchers until the 1980s. One of the first studies to explore the role that subtitles may play in second language acquisition was conducted by Lambert *et al.* (1981: 146), who, using standard subtitles with English as L1 (text) and French as L2 (audio), concluded that this combination was "unpromising" for improving comprehension and general language skills. Despite their unfavourable opinion, the paper is undisputedly pioneering in the field. Their results were soon partially contradicted by Price (1983) in a study with students at Harvard University in which she found that closed-captioned TV programmes helped students of English as a second language (ESL) better understand the video content. Since the publication of these early works, the increasingly pervasive nature of audiovisual materials and subtitles in our society has galvanised researchers' interest in exploring the potential that watching (and producing) subtitled audiovisuals can have in the honing of the various FLL skills.

The perceived ease with which results can be obtained, together with the instant impact on subjects, has lead vocabulary acquisition and recall to be two of the most frequently examined skills. Studies have been carried

out by experts in numerous language pairs, namely, Danan (1992) on English L1 and French L2, Karakaş and Sariçoban (2012) on Turkish L1 and English L2, Hsu *et al.* (2013) on Chinese L1 and English L2, Mousavi and Gholami (2014) on Persian L1 and English L2, and Birulés-Muntané and Soto-Faraco (2016) on Spanish L1 and English L2, Lertola (2019a) on English L1 and Italian L2, to name but a few. Listening comprehension is another skill that has received wide attention from scholars from the perspective of subtitling because audiovisuals promote listening by exposing students to the original speech contained in the soundtrack. Work on this topic include publications by scholars such as Markham (1989, 1999), Yoshino *et al.* (2000), Talaván (2009), Tsai (2010), Winke *et al.* (2010), Wang (2014) and Vanderplank (2016). Research has also been done on the acquisition of L2 speaking skills (Borrás and Lafayette, 1994; Arslanyilmaz and Pedersen, 2010; Jin, 2015), on the improvement of grammar (Van Lommel *et al.*, 2006), on reading comprehension skills (Guillory, 1998; Chen, 2012) and the potential that subtitles have in promoting incidental learning among viewers (Van de Poel and d'Ydewalle, 2001; Marzà and Torralba, 2015).

Research conducted on the topic has traditionally conceptualised the use of subtitles in the classroom from a passive perspective, in which the viewers' tasks are limited to listening to the soundtrack and reading the written subtitles. Such a premise has shaped the research conducted before the 2000s, and most of the research done afterwards, which, as illustrated in the above paragraph, tends to focus on a gamut of skills that may be enhanced thanks to the consumption of subtitles. In this respect, three types of subtitles have been typically explored from the perspective of second language acquisition: standard subtitles (from L2 speech to L1 text), reverse subtitling (from L1 speech to L2 text) and intralingual subtitling (from L2 speech to L2 text).

The technological advancements and digitalisation of audiovisual media seen in the late 1990s led to the development of specialist subtitling software, freely available on the Internet, that would have a substantial impact on the exploitation of subtitling in educational settings. Indeed, the affordances of technology brought about a more dynamic approach towards subtitling as students are now able to carry out new activities, such as creating their own subtitles or editing those produced by someone else, therefore increasing their level of engagement with the task. Worth mentioning in this respect are the EU-funded research projects LeVis (Learning via Subtitling, levis.cti.gr) and ClipFlair (Foreign Language Learning through Interactive Captioning and Revoicing of Clips, clipflair.net), the main objectives of which included the design and development of educational tools and activities intended for foreign-language teaching and learning (Sokoli, 2006, 2015). In the case of ClipFlair, an easy-to-navigate cloud-based platform was produced with the aim of facilitating the interaction between learners and technology, while at the same time allowing them to have a real hands-on experience in the production of subtitles.

The advantages of actively being involved in the creation of subtitles as compared to the passive experience of just watching them are numerous. In the first instance, producing their own subtitles gives students the feeling that "they play an active role in their own learning process" (Talaván, 2006: 316). Because of their markedly practical nature, subtitling tasks can be exploited in both formal educational contexts with the support of the teacher and in self-learning conditions. Additionally, subtitling as a process is alluring to students and tutors because it fosters three of the main traditional communicative skills, namely, listening, reading, and writing. It also invites learners to engage with the task in a more reflective manner as they will have to listen to the dialogue exchanges repeatedly and fully understand the text before they can proceed to the formulation of their subtitles. The honing of transferable skills, thanks to the learners' exposure to digital technology and different file formats (text, video, subtitles), as well as the practice of translation as a mediating and intercultural activity, also contribute to the educational added value of subtitling. One of the first scholars to take up this productive line of exploration is Talaván (2010, 2011), with other scholars, such as Lertola (2013, 2019b), following suit.

Researchers have found that producing standard subtitles helps with L2 listening comprehension skills (Williams and Thorne, 2000), incidental L2 vocabulary acquisition (Marzban and Zamanian, 2015) and the reinforcement of cultural awareness (Borghetti and Lertola, 2014). Reverse subtitling has been found to improve L2 writing skills (Talaván and Rodríguez-Arancón, 2014). Although the practice of intralingual subtitling has not been a popular topic in this line of research, as it does not involve translation, Lertola (2015, 2019b) suggests that it can be exploited to develop comprehension skills, rewording skills and summarising and paraphrasing abilities.

When it comes to the Chinese educational environment, the little research available on the value of subtitling for FLL focuses solely on the passive activity of reading subtitles. The experiment presented in the following pages is the first one to date to have been conducted with Chinese subjects on the more dynamic task of actually producing subtitles.

3. Methodology and materials

An experiment was conducted in November 2019, just before the pandemic, with 21 native Chinese speaking students from Harbin Engineering University, who were learning English as a second language. All participants were from non-English majors, with a science or engineering background. None of them had learnt a third language before the experiment and they did not have any previous experience in subtitling or translation. Although 21 participants started the experiment, it was completed by 17 subjects.

The experiment consisted of four sessions in total: a pre-test and initial questionnaire session, two activity sessions, and a post-test and final questionnaire session. In the first session, participants were given an English proficiency test consisting of a listening comprehension test with multiple-choice questions, a dictation, a vocabulary pre-test with sixty-one terms and a pre-test questionnaire with seven questions eliciting their background information as well as past experience in subtitling. Based on the result they obtained in the English proficiency test, students were allocated to one of the following three groups: (1) English to English subtitling group (EEG), 6 subjects, after one dropped out; (2) English to Chinese group (ECG), 7 subjects; and (3) control group (CG), 4 subjects, after three dropped out. To guarantee an evenly spread of students in terms of linguistic competence in the foreign language, so that the average English proficiency level would not suffer from striking differences across the different groups, students were selectively allocated to the groups: the student with the highest mark in the test (1) was allocated to group EEG, the one with the second highest mark (2) to group ECG and the participant with the third highest mark (3) to group CG. Then, the order of the group allocation was altered to CG (fourth highest mark), ECG (fifth highest score) and EEG (sixth highest result). The composition of the three groups is presented in Table 1:

Group EEG	Group ECG	Group CG
1	2	3
6	5	4
9	7	8
11	12	10
13	14	15
18	17	16
21	19	20
23	24	22

Table 1. Allocation of participants to groups.

The vocabulary test evaluated students' understanding of 61 words extracted from the video narrative which was to be used in the second and third sessions. It contained many medical terms that the students would not easily encounter in everyday life. Although they were prompted to list all the possible meanings for each of the lexical items, only the meaning intended in the video was considered to be correct when marking the test. One mark was given for a correct answer and zero for an incorrect or blank one, thus making a total maximum mark of sixty-one and a minimum one of zero.

Sessions 2 and 3 lasted two hours each and took place in consecutive weeks. No homework was given to students and all activities were done in class so that everyone had the same contact hours. The two experiment groups, EEG and ECG, were asked to perform a different subtitling task in each of the sessions, for which the cloud-based platform ClipFlair was used. The two clips used in the experiment were 5-minute-long extracts from popular science talks hosted by TED-Ed, namely, *How do glasses help us see* (for section 2) and *What would happen if you did not drink water* (for section 3). The two clips were narrated by the same male voice with an American accent. Before the second session, EEG and ECG were given instructions on how to use the platform. To facilitate the technical dimension of the tasks, previously timed templates with 77 subtitles for section 2 and 88 subtitles for section 3 were provided to the two subtitling groups. EEG participants were asked to subtitle, intralingually and verbatim, the dialogue in L2. For this task, they were provided with pre-timed subtitles in the same language, which contained only the initial letter of each caption line. ECG participants were given a file with verbatim L2 captions and asked to translate them into fluent L1. Meanwhile, the subjects in the control group, CG, performed summary-writing activities after having read the transcripts of the video, within the same time frame. All groups were allowed to access the internet and any dictionaries they deemed useful.

In the last session, a vocabulary post-test containing 61 words and a final questionnaire with a combination of three (CG) / four (EEG and ECG) closed questions and two (CG) / three (EEG and ECG) open-ended questions were administered to all participants, to elicit their opinion on the short experiment. The vocabulary post-test was the same as the pre-test, although the students did not know that it was the same, and the marking criteria applied were also the same for both tests, i.e., a potential minimum score of zero and a maximum of sixty-one. As already mentioned, the questionnaires used for the three groups were slightly different, but they all contained a combination of close-ended questions and attitudinal open-ended ones, in an attempt to reap the benefits of both approaches. Indeed, closed-ended questions were used in the hope that a higher response rate would be elicited from respondents, as they did not have to type too much. Subsequent analysis would be reasonably straightforward from a statistical perspective. One of the most important benefits of open-ended questions is the fact that they can potentially allow researchers to glean information that had not been anticipated, as respondents are given the opportunity to elaborate on their motivation, expectations and overall satisfaction.

An overview of all the various activities involved in the experiment, including their duration, is provided in Table 2:

Iterations	Activity	Duration	Groups
1	Initial test	2 hours	EEG, ECG and CG
2	Subtitling tasks	2 hours	EEG and ECG
2	Alternative language tasks	2 hours	CG
1	Final test	2 hours	EEG, ECG and CG

Table 2. Overview of experiment activities.

4. Results and discussion

According to the information compiled with the pre-test questionnaire, the participants were 19 to 26 year-old undergraduate and postgraduate students, who had been learning English as a foreign language for seven to 17 years prior to the experiment. None of the students had learnt a third language nor had they had any experience with subtitling before the experiment.

Numerical raw data collected during the experiment have been analysed using the SPSS statistical package v22, with the significance level set at $p<.05$ for all tests involved in the study. Given the small sample size of the experiment, non-parametric tests have been applied in the statistical analyses.

After the group allocation, a Kruskal-Wallis test performed in SPSS on the English proficiency test and the vocabulary pre-test of each of the three groups showed that the p was .941 and .272, respectively. Both values were greater than the set significance level of .05, indicating that the participants' marks did not show statistically significant differences among the groups and the distribution of participants was reasonable.

Figures 1, 2 and 3 show the scores attained by each of the participants on the vocabulary pre- and post-test, illustrated by group. The horizontal axis presents the individuals in the order of their performance on the pre-test, from the lowest score to the highest:
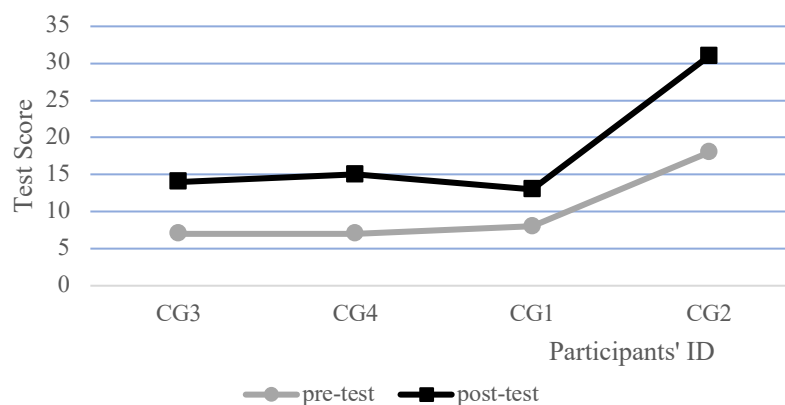

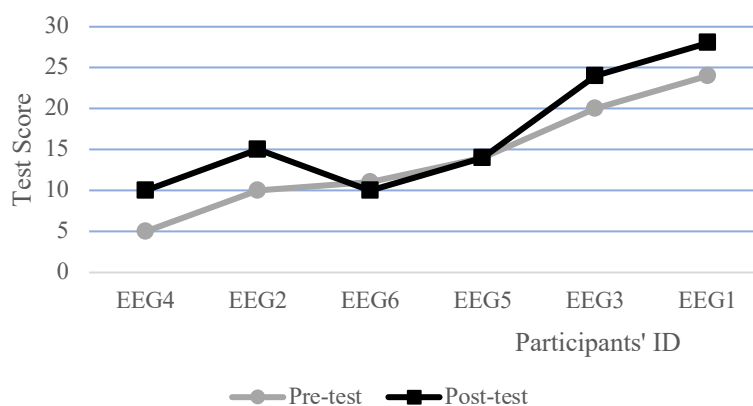*Figure 1. CG vocabulary pre- and post-test scores.*

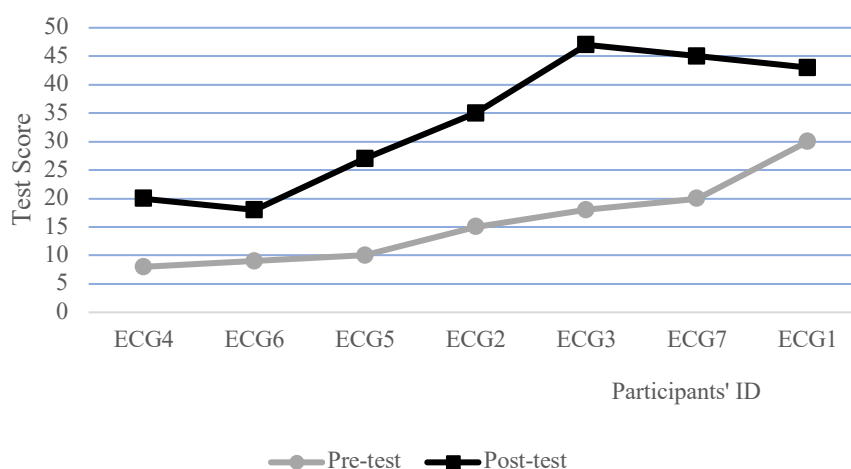*Figure 2. EEG vocabulary pre- and post-test scores.*

Figure 3. ECG vocabulary pre- and post-test scores.

As displayed in the figures, all participants scored a higher mark in the vocabulary post-test than in the pre-test, except for EEG5 and EEG6, where EEG5 obtained exactly the same result (14) before and after the subtitling activities and EEG6 scored one mark lower (10) after having done the subtitling activities. Theoretically, it is possible, albeit surprising, that the activities did not have a positive effect on students. In the case of EEG6, the participant gave a correct answer for the word 'stroke' in the pre-test but left it blank in the post-test. Interviewed after the experiment, the student confirmed that he was certain about the meaning of the term before the experiment. As the activity of his group was intralingual verbatim subtitling, participants came across words that sounded familiar, but whose spelling they were not sure about. In this particular case, the participant decided to use the online dictionary Hai Ci (dict.cn), which provides a fuzzy search function where the user types in the approximate spelling of a word and the dictionary then returns twenty fuzzy matched words according to the given spelling. The issue was therefore more related to the use of electronic resources than the actual subtitling. Indeed, the dictionary helped him produced more correctly spelt subtitles on the whole, but it also confused him by presenting so many similar words at the same time, which eventually led to his memory failure in the post-test.

To gain a contrastive view, the study analysed the variations in subjects' performance between pre- and post-test. As already mentioned, the scores of most participants improved and each individual's result is displayed in Table 3, together with the mean value and variance of each group:

Participants	CG	EEG	ECG
1	7	4	13
2	8	5	20
3	5	4	29
4	13	5	12
5		0	17
6		-1	9
7			25
Mean	8.25	2.8333	17.8571
Variance	8.6875	8.7083	45.2653

Table 3. Means and variances of improvements in the three groups.

To ascertain the potential variations across the results that could have been prompted by the different nature of the activities among the three groups, a Kruskal-Wallis test was first performed in SPSS. The result of the analysis showed that statistically significant variations (p=.002 <.05) in the improved scores exist among the three groups. However, since a Kruskal-Wallis test could not distinguish one group's variant from another, a further analysis using the Mann-Whitney U test was conducted, comparing two groups at a time.

First, the variance between CG and ECG was analysed. The significance run by the Mann-Whitney U test was *p*=.024 (<.05), suggesting that the improvement in the vocabulary test scores presented statistical significance between participants from these two groups. Given that the average score improvement of ECG (17.85) was more than the double of that of CG (8.25), it can be concluded that, by doing subtitling activities from L2 to L1, participants had acquired more words than those who had been doing traditional reading and writing activities.

The EEG and ECG data were then compared. The significance returned by the SPSS Mann-Whitney U test was *p*=.001 (<.05), proving that a significant difference also exists between these two other groups. Since the mean of the score improvement of ECG (17.85) was more than six times that of EEG (2.83), it can be concluded that performing interlingual subtitling activities from L2 to L1 is more efficient in vocabulary acquisition than doing intralingual subtitling activities from L2 to L2.

Finally, a comparison between CG and EEG was carried out. The significance generated by the SPSS Mann-Whitney U test was *p*=.019 (<.05), revealing that the improvements of CG and EEG also showed a significant difference. The mean improvement of CG (8.25) was nearly three times that of EEG (2.83), demonstrating that, in this study, carrying out traditional reading and writing activities resulted in more vocabulary acquisition than performing L2 intralingual subtitling activities.

The groups were also administered slightly different post-test questionnaires because of the variations in the course design and the data collected from the final questionnaires were not analysed in SPSS because answers from questionnaires were considered to be subjective and only objective data were analysed in SPSS. The first question asked students to rate the degree of difficulty of the reading materials (CG) or of the speech of the audiovisual materials (EEG and ECG) used in class, on a scale of 1 (very easy) to 5 (very difficult). The results are displayed in Figure 4:
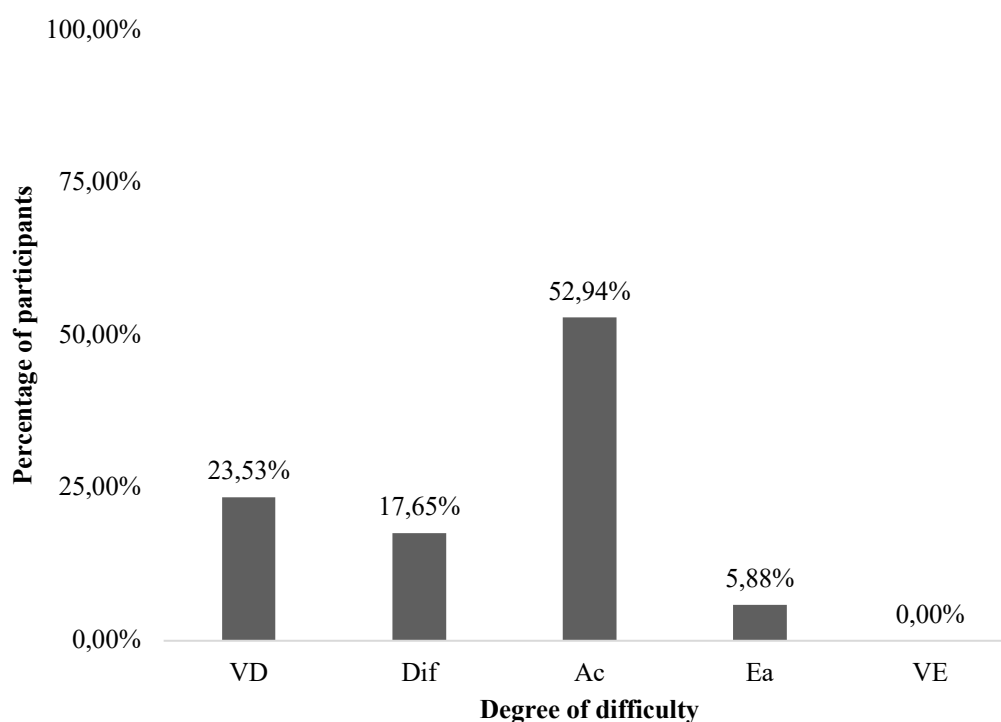
Figure 4. Participants' opinions on difficulty of source material.

In the participants' opinion, the difficulty of the materials used in the study was on the whole acceptable, leaning towards somewhat challenging. Indeed, more than half of them (9, 52.94%) reported that the difficulty of the materials was acceptable (Ac), while 41.17% found them to be difficult (Dif, 3, 17.65%) or very difficult (VD, 4, 23.53%). Only one student (5.88%) stated that the materials were easy (Ea) and no participants thought that they were very easy (VE).

The second question asked students their opinion as to whether they believed that they had acquired new vocabulary during the course. Two of the participants in CG reported having learnt new vocabulary while the other two did not. As for EEG, only one participant (16.67%) thought she had made improvements in vocabulary while the other five (83.33%) provided a negative answer, though three of them showed positive results on the vocabulary tests. In ECG the answers were more positive as four (57.14%) of the seven participants considered they had enriched their vocabulary. Interestingly, of the ten participants who reported no vocabulary acquisition, only two of them, both from EEG, obtained a negative or null result in the two vocabulary tests. The fact that eight students in total did not recognise the positive effects of the course on vocabulary acquisition, while their results point in the opposite direction, can be understood as an indication that the learning must have happened unconsciously.

The third question, which enquired about their interest in seeing this type of activity included in their formal English language classes, saw an enthusiastic 100% approval in EEG and ECG. Although participants from CG had not been exposed to subtitling in this study, three of them (75%) also showed an interest in incorporating these activities to their learning routines and only one of them (25%) recorded a negative predisposition. The data clearly show that students were overwhelmingly in favour of introducing variation to their learning and appreciated the novel and dynamic nature of subtitling.

The fourth question, addressed to EEG and ECG, aimed at teasing out the subjects' opinion about ClipFlair's virtual learning environment, by asking them to rate, from strongly agree (SAg) to strongly disagree (SD), the

statement, "the learning platform was convenient to use". The results obtained from the 13 subjects are illustrated in Figure 5:
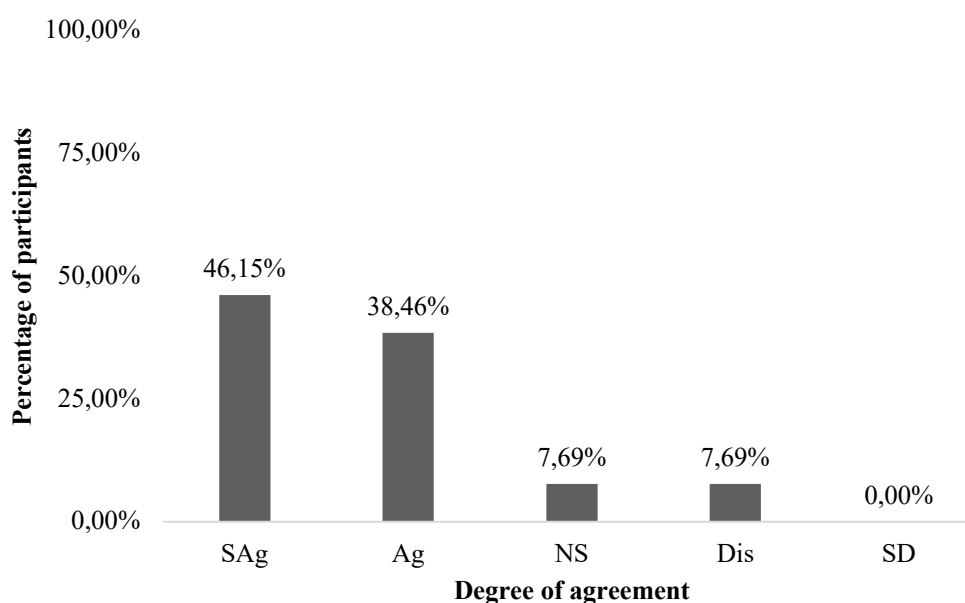

Figure 5. EEG and ECG's opinions on the convenience of ClipFlair.

Overall, their opinion was favourable (11, 84.61%), with six participants (46.15%) declaring that they 'strongly agreed' and 5 (38.46%) 'agreed' (Ag). One participant (7.69%) was non-committal with his 'not sure' (NS) response and only one student 'disagreed' with the statement provided. Some comments about technical glitches, raised in the open-ended questions, come to partially explain the unfavourable opinion of some of the participants and are discussed below.

In the post-test questionnaire administered to EEG and ECG, the first open-ended question asked them to describe in detail the technical issues, if any, that they had encountered while using ClipFlair. On the whole, students seemed to have found the process smooth and only two participants reported having experienced technical difficulties. ECG7 wrote that, while performing a task, the working window was accidentally amplified and she was unable to change the size back. This glitch resulted from the fact that the zoomable window in the platform allows users to adjust it into any size and the participant was not sufficiently familiar with ClipFlair to undo it. In the future, deactivating the edit option when designing the activity in the platform would solve the issue. She also reported a malfunction with the browser that caused all her work to disappear because she had not saved it beforehand. She complained that the file saving function in the platform was unnecessarily complicated as it requires users to specify a location every time they try to either save the activity or export the subtitle file, bearing in mind that they will need to replace any previous version of the file when saving the latest one. It is true to say that the system is rather unorthodox when it comes to the saving functionality; any future updates of the platform should consider a potential improvement in this area by, perhaps, allowing the creation of users' accounts so that past and ongoing projects can be saved online, in their respective personal accounts. EEG2 raised the issue that the subtitles were only editable when the video was paused, which is part of the platform design so that the subtitle window appears in synchrony with the subtitle being displayed on the video at that time. Although this function has the benefit of forcing users to stop and focus on a given subtitle, it makes it rather cumbersome if they need to listen to the same dialogue line more than once, which is a common occurrence when subtitling. Instead of having to manually move the cursor across the screen to edit the subtitle

and pause the video, a solution could be found whereby the video would pause automatically after, for instance, users have double-clicked the edit box to write their text. Alternatively, the video could continue to play but the subtitles would stop scrolling up when users are editing.

In the second open-ended question to all participants, they were asked to write a minimum of twenty Chinese characters to describe what they had gained from the experience and to record any opinion they had about the entire experimental course. In CG, students reported that the reading materials were rather long and contained challenging professional jargon from various medical fields, making the activities very demanding. However, some reflected that the activities were refreshing, as they had never done any summarising exercises in their L2 before, and effective in learning English. When describing their experiences, participants reported to have experienced progress in vocabulary acquisition as well as in their reading and summarising abilities in L2.

In the answers collected from EEG, only one participant volunteered feedback on the materials used in the course, noting that the speed of the narration in the clips was rather fast. The remaining comments pivoted around their involvement in the seminars. Five subjects felt that they had experienced significant progress in their listening comprehension skills, of which three revealed that they generally had difficulty when trying to recognise words that they had already learnt, which was the reason why they did not enjoy doing the listening tasks as the audio seemed more like a flow of desultory words and phrases to them than a set of distinct sentences carrying complete meanings. As a positive development, they felt that, particularly in the second subtitling activity, they were better at following the sentences heard in the dialogue and at making sense of the message being conveyed. One of the subjects highlighted the fact that she could distinguish words whose original pronunciation had been altered because of liaisons and elisions, while another, who claimed to have improved significantly, explained that, in the second activity, she had been able to provide more words in each of the subtitles than during the first activity, where she had only managed to distinguish and jot down one word for an entire caption in many cases. Two other participants, who did not mention having experienced any improvements, stated that the materials were too difficult for them, with one of them acknowledging that learning English through subtitling was a very interesting proposition because she could "acquire knowledge" in an "entertaining" way. Wisely taking advantage of the semiotic nature of audiovisual productions, one participant reported that she had learned to figure out words that she could not recognise from the audio thanks to the images, adding that she remained more focused and patient during the subtitling tasks than during other more traditional activities she had previously performed. Another participant reported that the four-hour activity gave her the feeling of being better at understanding fast-speaking audio and that she now felt "relieved and confident" when doing other listening activities. Interestingly, all seven students in this group mentioned that the transcribing and subtitling tasks were rather exhausting though three of them conceded that they had become less so during the second activity because the first task had allowed them to become familiarised with the activity and the materials.

ECG participants also provided positive feedback on the course, describing it as a more acceptable, unique, and entertaining way of learning English than some of the other tasks they were accustomed to. Some of them reported having made progress not only in vocabulary acquisition but, because of the use of templates, also in reading comprehension. Nevertheless, two of the participants expressed their concern as to whether the progress they made during the subtitling activities was worth the time and effort they had invested in them. In this regard, they felt that the extent of their progress was too small when compared to the vast amount of time and energy they had put into the various activities, although two participants admitted that the activities may be effective in the long run as their learning might have become more systematic.

The third and final open-ended question of the post-test questionnaire asked them to give their opinion on the two activity sessions. All four CG participants wished that the teacher had involved more lecturing in the classes

rather than letting them do the activities alone and providing support occasionally. EEG participants provided several suggestions on how to redesign the course, including selecting videos with less demanding content and with a slower speaking speed, lengthening the course to include some more sessions, and discussing the meaning of difficult words before the activity. One participant proposed that the ClipFlair platform should have an auto-save function so that any unexpected glitches would not result in the need to retype everything. Finally, ECG reported that doing subtitling activities on the basis of a verbatim English template risks disregarding the audio component of the video; a shortcoming that could be overcome in the future by resorting to templates containing blanks in certain sections or, depending on the technical know-how of the students, by requesting them to produce the spotting, i.e., deciding the in and out times of the subtitles, and translation of the subtitles. Two other comments provided by the students underlined the facts that the sessions could have involved more interaction with the teacher and that subtitling from Chinese audio to English (reversed subtitling) could have also been practised. Both suggestions can easily be taken on board in further experiment designs but the controlled variables can be difficult to manage when close interaction between the teacher and the students is involved, while doing reverse subtitling would primarily enhance L2 writing skills rather than listening and reading comprehension skills.

5. Conclusion

Through a short course that lasted for four two-hour sessions, the main purpose of this study was to discover and compare the English vocabulary acquisition results obtained by three groups of Chinese participants when carrying out intralingual or interlingual subtitling activities or more traditional ones like reading, summarising and translating written texts.

In the pre-test session, neither the scores of the English proficiency test nor those of the vocabulary test showed significant differences between participants from different groups, suggesting that the allocation of groups explained in section 3 was rational. After conducting the various FL activities over the course of two weeks, participants were administered a vocabulary post-test, which was exactly the same as the pre-test, in order to evaluate any potential improvements in their performance by comparing the post-test scores with the ones obtained in the pre-test. Each of the groups' performance showed significant variance with the other two groups when Whitney U tests were conducted in SPSS, with ECG displaying the greatest improvements (17.85 on average), EEG the lowest (2.83), and CG in the middle (8.25). In answer to the initial research question, and rather unsurprisingly, the results confirm that all three scenarios favoured the acquisition of L2 vocabulary. Yet, of the three conditions, conducting interlingual subtitling activities in the standard subtitling combination (i.e., from L2 foreign language speech into L1 mother tongue text) is decidedly the most advantageous. Another interesting finding resides in the fact that the impact of doing intralingual subtitling, at least in this experiment, is not as significant as in the non-subtitling condition modelled on more traditional L2 education tasks. Given that one of the main differences in design between ECG / CG and EEG is the practice of translation tasks required in the first two conditions, it can be argued that translation is indeed a productive tool in FLL, thus dynamiting the precepts of certain educational paradigms, such as the communicative approach, whereby emphasis should be placed on the exclusive use of the foreign language in the classroom, which led Duff (1989: 5) to decry: "translation has been generally out of favour with the language teaching community. (Almost, we might say, 'sent to Siberia'!)".

Participants' feedback was, on the whole, very positive and constructive. Despite thinking that the materials selected were slightly challenging, they enjoyed the experience and were willing to see subtitling activities

featuring more prominently in higher education L2 classes. Students in the L2 to L1 subtitling group preferred this type of activity when learning new vocabulary over more traditional approaches, though some of them felt that there was an imbalance, since the amount of effort dedicated to the tasks seemed to be too onerous when compared to the benefits they could reap. In this respect, longitudinal experiments should be conducted to evaluate whether the perceived 'extra effort' invested in these activities pays its dividends in the long run, in the form of longer memory retention of the terms learnt in class.

Although the participants approved of the subtitling and language learning platform, ClipFlair, they also put forward suggestions that they thought would improve its functionality, particularly in the way in which their work can be saved. Given the fact that the funding for the project ran out after its completion, it is unlikely that a revamped version of ClipFlair will see the light. Using desktop subtitling software in the FLL classroom as an alternative runs the risk of being too complex from a logistical perspective but the great advances witnessed in the development of cloud-based platforms, both in the areas of subtitling and dubbing, are full of promise (Bolaños-García-Escribano and Díaz Cintas, 2020).

In terms of the participants' self-assessment of their performance and perceived improvement in the vocabulary tests, a mismatch has been detected between their opinion and the actual results. Indeed, despite generating the greatest statistical improvement, ECG did not give much credit to their progress, intimating that vocabulary acquisition might take place in an unconscious manner. EEG thought highly of the subtitling activities as a means to improving their English proficiency though the overall results of this group were the more modest in terms of lexical expansion. An interpretation of this circumstance could be that the seminars were instrumental in boosting students' confidence in their FL abilities and that the intralingual subtitling activities may well have resulted in improvement of other skills that were not tested in the experiment, such as L2 listening comprehension. A similar finding derived from the comments provided in the open-ended questions of the final questionnaire by some CG and ECG participants, who felt that the activities had helped them enhance their reading comprehension skills. Yet again, this is a rather subjective impression that was not tested or quantified in the experiment and that opens the doors to further research on the different language skills that can be honed when practising activities of this nature.

The results obtained in this empirical study show that, when compared to traditional methods of learning an L2 or to intralingual subtitling in L2, activities based on standard interlingual subtitling from L2 to L1 are the most effective in enhancing vocabulary acquisition of the foreign language. Nonetheless, and despite the positive results returned by the experiment, some aspects of the study could be amended in future repetitions. First, the sample size in the study was rather small and future iterations should be conducted with more subjects to boost the chances of drawing more reliable conclusions. Second, the experiment should last for a longer period so that the long-term effects of subtitling on FL learning can be ascertained. Third, to overcome the shortcoming that when working with a verbatim template in L2 the soundtrack can be disregarded, activities could be designed for students to work with an incomplete template, to fill in the missing blanks, or for them to take care of the spotting too. In both scenarios, the results will help researchers gain a more complete picture of the potential offered by the multiple options of exploiting subtitling. Finally, biometric sensors like eye trackers as well as screen recorders could be brought into the experimentation as means that would not only help triangulate the results but also have the potential of shedding further light into the cognitive processes scaffolding the practice of subtitling.

As highlighted at the beginning of this paper, the use of subtitles, and especially didactic subtitling (Talaván, 2010), as a way of diversifying FL teaching and learning has become increasingly popular in our digital society. Its positive impact on students' academic performance together with the fascinating technological developments

currently taking place in the area of cloud subtitling (and dubbing), which help lower the technical complexity around these tasks, herald a most promising future for these audiovisual activities in the foreign language classroom, which are bound to gain prominence and visibility in the years to come.

Declaration of conflicting interests

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding

The author(s) received no financial support for the research, authorship, and/or publication of this article.

About the authors

Jorge Díaz Cintas is Professor of Translation Studies and founding director (2013-2016) of the Centre for Translation Studies (CenTraS) at University College London. He is the author of numerous articles, special issues and books on audiovisual translation. He is the Chief Editor of the Peter Lang series *New Trends in Translation Studies* and the recipient of the Jan Ivarsson Award (ESIST, 2014) and the Xènia Martínez Award (ATRAE, 2015) for invaluable services to the field of audiovisual translation.

Chengcheng Wang is a PhD candidate at University College London, where she is researching the didactic role of subtitling in the foreign language classroom. She works as a freelance subtitler and teaches subtitling modules at various universities, both in China and in the United Kingdom.

References

Arslanyilmaz, A. & Pedersen, S. (2010). Improving language production using subtitled similar task videos. *Language Teaching Research, 14*(4), 377-395.

Birulés-Muntané, J. & Soto-Faraco, S. (2016). "Watching subtitled films can help learning foreign languages". *PloS one* 11(6): e0158409.

Bolaños-García-Escribano, A. & Díaz Cintas, J. (2020). The cloud turn in audiovisual translation. In Ł. Bogucki & M. Deckert (Eds), *The Palgrave Handbook of Audiovisual Translation and Media Accessibility* (pp. 519-544). Basingstoke: Palgrave Macmillan.

Borghetti, C. & Lertola, J. (2014). Interlingual subtitling for intercultural language education: a case study. *Language and Intercultural Communication*, *14*(4), 423-440.

Borrás, I. & Lafayette, R. C. (1994). Effects of multimedia courseware subtitling on the speaking performance of college students of French. *The Modern Language Journal*, *78*(1), 61-75.

Brett, D. & González-Lloret, M. (2009). Technology-enhanced materials. In M. H. Long & C. J. Doughty (Eds), *The Handbook of Language Teaching* (pp. 351-369). Chichester: Wiley-Blackwell.

Chen, M.-L. (2012). Effects of the order of reading text or viewing a film and L1/L2 captions on reading comprehension. *Perceptual & Motor Skills: Learning & Memory, 115*(1), 18-26.

Danan, M. (1992). Reversed subtitling and dual coding theory: new directions for foreign language instruction. *Language Learning, 42*(4), 497-527.

Danan, M. (2015). Subtitling as a language learning tool: past findings, current applications, and future paths. In Y. Gambier, A. Caimi & C. Mariotti (Eds), *Subtitles and Language Learning: Principles, Strategies and Practical Experiences* (pp. 41-61). Bern: Peter Lang.

Díaz Cintas, J. & Fernández Cruz, M. (2008). Using subtitled video materials for foreign language instruction. In J. Díaz Cintas (Ed.), *The Didactics of Audiovisual Translation* (pp. 201-214). Amsterdam: John Benjamins.

Duff, A. (1989). *Translation*. Oxford: Oxford University Press.

d'Ydewalle, G. & Van de Poel, M. (1999). Incidental foreign-language acquisition by children watching subtitled television programs. *Journal of Psycholinguistic Research, 28*(3), 227-244.

Gambier, Y. (2015). Subtitles and language learning (SLL): theoretical background. In Y. Gambier, A. Caimi & C. Mariotti (Eds), *Subtitles and Language Learning: Principles, Strategies and Practical Experiences* (pp. 63-82). Bern: Peter Lang.

Guillory, H. G. (1998). The effects of keyword captions to authentic French video on learner comprehension. *CALICO Journal, 15*(1-3), 89-109.

He, X. 2015. *Tapping the Pedagogical Potential of Subtitling in Foreign Language Learning*. MA dissertation. London: University College London.

Hsu, C., Hwang, G.-J., Chang, Y.-T. & Chang, C.-K. (2013). Effects of video caption modes on English listening comprehension and vocabulary acquisition using handheld devices. *Educational Technology & Society, 16*(1), 403-414.

Incalcaterra McLoughlin, L. & Lertola, J. (2011). Learning through subtitling: subtitling as an aid to language learning. In L. Incalcaterra McLoughlin, M. Biscio & M. Á. Ní Mhainnín (Eds), *Audiovisual Translation: Subtitles and Subtitling: Theory and Practice* (pp. 243-263). Oxford: Peter Lang.

Jin, S. (2015). 基于提高高中生英语口语流利性的视频素材应用研究 [*The Application of Video Materials in Improving Senior High School Students' Oral Speaking Fluency*]. MA dissertation. Harbin: Harbin Normal University.

Karakaş, A. & Sariçoban, A. (2012). The impact of watching subtitled animated cartoons on incidental vocabulary learning of ELT students. *Teaching English with Technology, 12*(4), 3-15.

Lambert, W. E., Boehler, I. & Sidoti, N. (1981). Choosing the languages of subtitles and spoken dialogues for media presentations: implications for second language education. *Applied Psycholinguistics, 2*, 133-148.

Lertola, J. (2013). *Subtitling New Media: Audiovisual Translation and Second Language Vocabulary Acquisition*. PhD thesis. Galway: National University of Ireland.

Lertola, J. (2015). Subtitling in language teaching: suggestions for language teachers. In Y. Gambier, A. Caimi & C. Mariotti (Eds) *Subtitles and Language Learning: Principles, Strategies and Practical Experiences* (pp. 245-268). Bern: Peter Lang.

Lertola, J. (2019a). Second language vocabulary learning through subtiltign. *Spanish Journal of Applied Linguistics*, *32*(2), 486-514.

Lertola, J. (2019b). *Audiovisual Translation in the Foreign Language Classroom: Applications in the Teaching of English and Other Foreign Languages*. Research-publishing.net.

Markham, P. L. (1989). The effects of captioned television videotapes on the listening comprehension of beginning, intermediate and advanced ESL students. *Educational Technology, 29*(10), 38-41.

Markham, P. L. (1999). Captioned videotapes and second-language listening word recognition. *Foreign Language Annals, 32*(3), 321-238.

Marzà, A. & Torralba, G. (2015). Incidental language learning through subtitled cartoons: is it possible in a dubbing country? In Y. Gambier, A. Caimi & C. Mariotti (Eds) *Subtitles and Language Learning: Principles, Strategies and Practical Experiences* (pp. 199-220). Bern: Peter Lang.

Marzban, A. & Zamanian, M. (2015). The impact of the subtitling task on vocabulary learning of Iranian EFL learners. *Journal of Applied Linguistics and Language Research, 2*(1), 1-9.

Mousavi, F. & Gholami, J. (2014). Effects of watching flash stories with or without subtitle and reading subtitles on incidental vocabulary acquisition. *Procedia-Social and Behavioural Sciences, 98*, 1273-1281.

Neves, J. (2004). Language awareness through training in subtitling. In P. Orero (Ed.) *Topics in Audiovisual Translation* (pp. 127-140). Amsterdam: John Benjamins.

Paivio, A. (1986). *Mental Representations: A Dual Coding Approach*. New York: Oxford University Press.

Price, K. (1983). Closed-captioned TV: an untapped resource. *MATSOL Newsletter, 12*, 1-8.

Sokoli, S. (2006). Learning via Subtitling (LvS). A tool for the creation of foreign language learning activities based on film subtitling. In *MuTra Proceedings 2006 – Multidimensional Translation: Audiovisual Translation Scenarios*. Copenhagen: University of Copenhagen. <http://www.sub2learn.ie/downloads/2006_sokoli_stravoula.pdf> [22/08/2021].

Sokoli, S. (2015). ClipFlair: foreign language learning through interactive revoicing and captioning of clips. In Y. Gambier, A. Caimi & C. Mariotti (Eds), *Subtitles and Language Learning: Principles, Strategies and Practical Experiences* (pp. 127-147). Bern: Peter Lang.

Sokoli, S., Zabalbeascoa, P. & Fountana, M. (2011). Subtitling activities for foreign language learning: what learners and teachers think. In L. Incalcaterra Mcloughlin, M. Biscio & M. Á. Ní Mhainnín (Eds), *Audiovisual Translation: Subtitles and Subtitling: Theory and Practice* (pp. 219-242). Oxford: Peter Lang.

Talaván, N. (2006). Using subtitles to enhance foreign language learning. *Porta Linguarum, 6*, 41-52.

Talaván, N. (2009). *Aplicaciones de la traducción audiovisual para mejorar la comprensión oral de inglés*. PhD Dissertation. Madrid: Universidad Nacional de Educación a Distancia.

Talaván, N. (2010). Subtitling as a task and subtitles as support: pedagogical applications. In J. Díaz Cintas, A. Matamala & J. Neves (Eds), *New Insights into Audiovisual Translation and Media Accessibility* (pp. 285-299). Amsterdam: Rodopi.

Talaván, N. (2011). A quasi-experimental research project on subtitling and foreign language acquisition. In L. Incalcaterra-McLoughlin, M. Biscio & M. Á. Ní Mhainnín (Eds), *Audiovisual Translation Subtitles and Subtitling. Theory and Practice* (pp. 197-217). Oxford: Peter Lang.

Talaván, N. & Rodríguez-Arancón, P. (2014). The use of reverse subtitling as an online collaborative language learning tool. *The Interpreter and Translator Trainer, 8*(1), 84-101.

Tsai, F.-H. (2010). *Integrating Feature Films with Subtitles to Enhance the Listening Comprehension of Students Attending College in Taiwan*. PhD thesis. San Diego: Alliant International University.

Van de Poel, M. & d'Ydewalle, G. (2001). Incidental foreign-language acquisition by children watching subtitled television programs. In Y. Gambier & H. Gottlieb (Eds), *(Multi)Media Translation: Concepts, Practices and Research* (pp. 259-274). Amsterdam: John Benjamins.

Vanderplank, R. (1988). The value of teletext sub-titles in language learning. *ELT Journal, 42*(4), 272-281.

Vanderplank, R. (1990). Paying attention to the words: practical and theoretical problems in watching television programmes with uni-lingual (CEEFAX) subtitles. *System, 18*(2), 211-234.

Vanderplank, R. (2015). Thirty years of research into captions/same language subtitles and second/foreign language learning: distinguish between 'effects of' subtitles and 'effects with' subtitles for future research. In Y. Gambier, A. Caimi & C. Mariotti (Eds), *Subtitles and Language Learning: Principles, Strategies and Practical Experiences* (pp. 19-40). Bern: Peter Lang.

Vanderplank, R. (2016). *Captioned Media in Foreign Language Learning and Teaching: Subtitles for the Deaf and Hard-of-Hearing as Tools for Language Learning*. Basingstoke: Palgrave Macmillan.

Van Lommel, S., Laenen, A. & d'Ydewalle, G. (2006). Foreign-grammar acquisition while watching subtitled television programmes. *British Journal of Educational Psychology, 76*, 243-258.

Wang, Y. (2014). *The Effects of L1/L2 Subtitled American TV Series on Chinese EFL Students' Listening Comprehension*. MA dissertation. Lansing: Michigan State University.

Williams, H. & Thorne, D. (2000). The value of teletext subtitling as a medium for language learning. *System, 28*(2), 217-228.

Winke, P., Gass, S. & Sydorenko, T. (2010). The effects of captioning videos used for foreign language listening activities. *Language Learning & Technology, 14*(1), 65-86.

Yoshino, S., Kano, N. & Akahori, K. (2000). The effect of English and Japanese captions on the listening comprehension of Japanese EFL students. *Language Laboratory, 37*, 111-129.