UNIVERSITAT DE BARCELONA

Grau d'Estudis Anglesos

Treball de Fi de Grau
Curs 2022-2023

Examining the Effectiveness of Captions in EFL Slang Terms and Idioms

Estudiant: Maria Port Puig

**Tutora:** Dr. Maria Luz Celaya Villanueva

Barcelona, 16 de Juny de 2023

Declaració d'autoria

Amb aquest escrit declaro que soc l'autor/autora original d'aquest treball i que no he emprat per a la seva elaboració cap altra font, incloses fonts d'Internet i altres mitjans electrònics, a part de les indicades. En el treball he assenyalat com a tals totes les citacions, literals o de contingut, que procedeixen d'altres obres. Tinc coneixement que d'altra manera, i segons el que s'indica a l'article 18 del capítol 5 de les Normes reguladores de l'avaluació i la qualificació dels aprenentatges de la UB, l'avaluació comporta la qualificació de "Suspens".

Barcelona, a 16 de juny de 2023

Signatura:
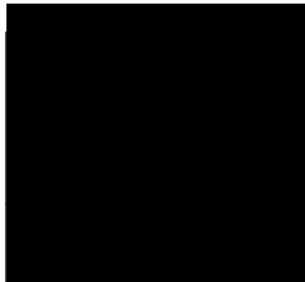
ACKNOWLEDGEMENTS

ABSTRACT

The use of audiovisual materials as a language acquisition resource has greatly increased over the last years as they are an engaging, multimodal form of input. As an effective tool for L2 learning, the use of on-screen text further encourages the acquisition of language, especially vocabulary. The present study, therefore, aims to examine the effectiveness of captions for the acquisition of British vocabulary, more specifically, slang terms and idioms. Two groups (n=6) of proficient-level students of English were exposed to a 5-minute video with captions and without captions, respectively, and were asked to take a test immediately before and after the viewing. Findings show that both groups improved their performance after watching the excerpt, regardless of the condition, although the captioned group scored a higher percentage of answers in the post-test. Thus, results demonstrate that captions foster the acquisition of English vocabulary, as suggested by previous studies.

Keywords: Acquisition of L2 English, audiovisual materials, captions, vocabulary learning.

RESUM

L'ús de materials audiovisuals com a recurs d'adquisició d'una segona llengua ha augmentat considerablement al llarg dels últims anys, ja que aquests ofereixen un *input* multimodal i atractiu. Com a eina efectiva per a l'aprenentatge d'una segona llengua, l'ús de text en pantalla afavoreix l'adquisició del llenguatge, especialment del vocabulari. Aquest estudi, per tant, té com a objectiu analitzar l'eficàcia dels subtítols en l'adquisició de vocabulari britànic, més concretament, argot i locucions. Dos grups (n=6) d'estudiants amb un nivell competent d'anglès van ser exposats amb subtítols i sense subtítols, respectivament, a un vídeo de 5 minuts i, immediatament abans i després de la visualització, se'ls va demanar que féssin un test. Tot i que els resultats mostren que ambdós grups van millorar el seu rendiment després de veure el vídeo, independentment de la condició, el grup amb subtítols va encertar un percentatge de respostes més alt en el posttest. Per tant, els resultats demostren que els subtítols fomenten l'adquisició del vocabulari anglès, tal i com es suggereix en estudis previs.

Paraules clau: Adquisició de l'anglès com a segona llengua, materials audiovisuals, subtítols, aprenentatge de vocabulari.

TABLE OF CONTENTS

1 INTRODUCTION

Over the last decades, technology has increasingly been introduced to people's daily lives. In fact, the immediate availability to audiovisual materials, such as films, documentaries and series, among others, has piqued people's interest for the multimodal characteristic of human communication (Jewitt, 2013, cited in Muñoz, 2022). The utilization of multimedia, however, has recently been considered as a learning tool. Audiovisual materials present the advantage of being entertaining resources whose access is unconstrained both temporally and spatially, and that allow exposure to a wide variety of linguistic features; hence, seeing that students can benefit from multimedia both inside and outside the classroom context, researchers such as Frumuselu et al. (2015) have considered it a means to bridge the gap between intentional and incidental learning.

Nonetheless, notwithstanding the value of audiovisual materials on their own, numerous studies have proven throughout the years that the addition of on-screen text aids in the acquisition of language. Yet, there are certain aspects such as level of proficiency that should be taken into account in order to determine whether students will benefit more from subtitles or captions; in fact, most researchers agree on the fact that the latter are essential for proficient learners of a language, despite them being "originally developed for hearing-impaired" audiences (Montero Perez et al., 2013). Actually, researchers have analysed the effects of different types of captions for L2 learning, but whereas some have claimed that variations are no more useful than plain captions, others have straightforwardly stated that, as a matter of fact, students benefit equally from the captioned and non-captioned conditions.

Thus, the aim of this paper is, on the one hand, to review research on the benefits of audiovisual materials materials for L2 learning and, on the other hand, to demonstrate the effectiveness of captions for the acquisition of English vocabulary in proficient-level students. In order to do so, a study on captioned and non-captioned conditions will be conducted and its data will be subsequently analysed, leading to a discussion on the findings that will determine the value of on-screen text transcriptions for language learning. To put it another way, assessing the results obtained from a test taken by a group (n=6) of proficient-level students of English who will have to watch a video excerpt captioned and non-captioned, the effectiveness of either condition for the acquisition of British slang terms and idioms will be examined.

2 REVIEW OF THE LITERATURE

2.1 Definition of audiovisual materials

Audiovisual materials are those resources that present audio and visual input, such as films, sitcoms, and TV programmes (Frumuselu et al., 2015), and whose popularity rises due to both their learning and entertaining characteristics (Bruti, 2016); that is, audiovisual materials are an engaging, "multimodal form of input" (Muñoz, 2022, p.125) that not only fosters L2 comprehensibility through the simultaneous use of sound, image and text, but also can be accessed anytime anywhere (Collins & Muñoz, 2016, cited in Muñoz, 2022). As Bruti (2016, p.191) explains, "scriptwriters select specific features of speech that are widely accepted and recognized as such by the audience", aiming to create a more spontaneous, native-like dialogue. Hence, seeing that scripted dialogues contain a high amount of non-standard language (Montero Perez, 2022), viewers are exposed to true-to-life, colloquial language when met with an audiovisual excerpt (Abrams, 2014, & Frumuselu et al., 2015). Nonetheless, between television and the cinema, TV series include less homogeneous language than films because of their shorter time limit; to put it another way, series have a more varied language in order to create natural, short dialogues as well as include plot twists to keep the audience on tenterhooks (Bruti, 2016). The degree of natural language present in scripted language, however, may depend on the genre, as pointed out by Montero Perez (2022); for instance, sitcoms present more natural conversations than soap operas.

2.2 Audiovisual materials in the classroom

According to Montero Perez et al. (2013), the use of multimedia has greatly increased over the last years as a language acquisition resource, seeing that they are perceived as a technique that fosters learning, including skills such as "listening comprehension, vocabulary, and grammar" (Barón, & Celaya, 2022, p.2); that is, it may benefit learners by improving their listening ability and vocabulary acquisition, as well as their oral proficiency (Montero Perez et al., 2013). In fact, audiovisual materials allow students to acquire contextualized L2 vocabulary and comprehend its meaning through both verbal and non-verbal input (Elgort et al., 2018, & Lin and Siyanova-Chanturia, 2015, cited in Montero Perez, 2022), seeing that, as multimodal resources, the simultaneous activation of aural, visual and written inputs allow audiovisual materials to become useful learning resources (Montero Perez, 2022).

Bruti (2016) emphasises the importance of selecting excerpts or clips rather than watching a whole film or episode seeing that some learners might find the latter more difficult to follow and understand, especially when the dialogues or the language are not understood; the researcher claims that short excerpts, on the contrary, allow students to focus on each component of the input despite

being a decontextualized audiovisual material; in other words, focusing on specific scenes or clips allows to analyse and understand the language better because learners are able to manage less input at once (Kaiser, 2011). In order to facilitate the teachers' lesson preparation, Kaiser (2011, p.236) states that the Berkeley Language Center has created a Library of Foreign Language Film Clips (LFLFC), an "Internet-based database of film clips" in which films are reduced to clips "no longer than four minutes" for "pedagogical reasons". However, Kaiser (2011) points out the importance of the teacher providing context to students before watching an excerpt, seeing that decontextualization may lead to meaning being lost, especially taking into account that some conversations may talk about previous scenes.

As regards L2 learning, Pemberton, Fallahkhair, and Masthoff (2004) explain that, whereas intentional or formal learning takes place in the classroom, incidental or informal learning takes place outside the classroom and is unconsciously acquired; hence, audiovisual materials do not only motivate learners, but also encourage incidental language acquisition, allowing them not only to learn inside, but also outside the classroom (Frumuselu et al., 2015). In other words, multimedia resources allow to bridge the gap between intentional and incidental learning when used as a learning tool (Frumuselu et al., 2015). However, Sockett (2012, cited in Montero Perez, 2022) points out the importance of L2 teachers' role in fostering learners to use audiovisual materials as an incidental learning tool; that is, L2 teachers should introduce video excerpts with on-screen text in the classroom to encourage learners to watch them autonomously. Thus, Montero Perez (2022) suggests three key factors should be taken into account in order to achieve this: selecting the audiovisual excerpts according to the student's learning goal, demonstrating the benefits that subtitles and captions have for L2 learners, and finally introducing students to the different types of written support available, as well as the advantages and disadvantages that each has. Barón and Celaya (2022) point out, however, that contextualized audiovisual excerpts encourage L2 learning regardless of textual support, that is, captions and subtitles (see section 2.3). The reason behind this could be the fact that multimedia is not artificially created or manipulated to suit L2 learners needs, but it is rather addressed to native speakers (Frumuselu et al., 2015).

2.3 Subtitles and captions for L2 acquisition

Audiovisual translation mainly comprises interlingual subtitling or subtitles and intralingual subtitling or captions. Suárez (2018, p.3) describes the first as an on-screen text translation of a video's aural input from a foreign language into the mother language, whereas the latter is an on-screen text transcription of a video's aural input in the same language that usually serves as an aid for "deaf and hearing-impaired audiences". Despite their original purpose, Montero et al. (2013)

mention that the numerous benefits of subtitles and captions encouraged their introduction in the classroom. However, some similarities and differences should be taken into account in order to decide whether to use subtitles or captions, as explained in what follows.

Mayer (2009, cited in Frumuselu et al., 2015) claims in his *Multimedia Learning Theory* that learners benefit from textual support, whether it be subtitles or captions (see line 5 in section 2.1), when met with an audiovisual excerpt as they are more prone to internalize the language or linguistic concepts; that is, students being exposed not only to visual, but also to oral and written native input may encourage their language acquisition (Frumuselu et al., 2014). As regards audiovisual materials' multimodality, Sweller (2005, cited in Frumuselu et al., 2015) suggests in his *Theory of Cognitive Load* that learners definitely benefit from being exposed to different channels at the same time, as they work together and complete one another to avoid cognitive overload.

According to Bruti (2016), both interlingual subtitles and intralingual captions do not only allow learners to check whether they have understood the input correctly by being able to compare the audio with the text, but they also reinforce the students' vocabulary retention; to put it another way, subtitles and captions allow to "bridge the […] gap between the development of skills in reading comprehension and listening comprehension" (Montero Perez et al., 2013, p.731). Likewise, Vanderplank (1988, p.280), points out the "mediating" characteristic of both subtitles and captions as they assist learners and make up for their low listening skills (Montero Perez et al., 2013); that is, written support can be considered a "means to outweigh deficient vocabulary size and overcome listening problems" (Montero Perez et al., 2013, pp.721-722). Notwithstanding the benefits of multimedia on its own, Frumuselu et al. (2015) point out that being able to read the conversation of a specific excerpt while listening to it in the L1 fosters language and vocabulary acquisition, seeing that it allows learners to reinforce what they hear. Besides, written support encourages learners to consciously focus on the form of new expressions and other input, allowing them to create a relationship between form and meaning, as well as isolate words in order to comprehend them (Montero Perez et al., 2013).

Nonetheless, as suggested in Vanderplank (1988)'s study, the effectiveness of subtitles and captions may vary from learners, especially as regards their proficiency level. Whereas beginner-level students can benefit more from subtitles, mainly because the text becomes a visual support that makes up for their lack of L2 knowledge, post-intermediate-level students can benefit more from captions because their L2 fluency allows them to focus on vocabulary recognition, language acquisition and comprehension of the text (Kaiser, 2011, p.244, & Montero Perez et al., 2013, p.733). Krashen (1985, p.3) supports Vanderplank's idea by suggesting that advanced learners would benefit more from captions seeing that their knowledge of the L2 allows them to take the

input in, whereas beginner levels would benefit more from subtitles because their lack of "self-confidence" as regards their L2 knowledge leads to them blocking the L2 input and, inevitably, failing in language acquisition. As Pujadas & Muñoz (2019, p.178, cited in Montero Perez, 2022) put it, captions may increase the students' self-assurance on their language skills by allowing a "more relaxed viewing", seeing that L2 words can be easily distinguished when they are orthographically represented (Majuddin et al., 2021). However, although subtitles allow indeed learners to process and comprehend better the input, leading to a deeper acquisition as they are able to remember the input over time, captions may become a mere textual support for proficient L2 learners, as captions simply allow them to visualize the audio of the excerpt (Frumuselu et al., 2015). Captions, nonetheless, can be essential for beginner learners as they may reinforce their low-level listening comprehension (Montero Perez et al., 2013).

Majuddin et al. (2021), on the contrary, argue that the benefits of captions are not determined by learner's proficiency level but rather by the modality of the post-test; that is, participants in the captioned group performed significantly better on comprehension tests than in productive tests. Likewise, Montero Perez (2022) states that, whereas captions are a mere transcription of the L2 input, subtitles provide its translated account, allowing learners to improve their L2 comprehension regardless of their proficiency level.

2.4 Effectiveness of captions in audiovisual materials

Notwithstanding the benefits of captions for L2 learning (see section 2.3), Montero Perez (2022) mentions in her article that its effectiveness might be conditioned not only by the learners' viewing conditions, but also by the techniques applied to the on-screen text, as will be seen below.

2.4.1 Learner's viewing conditions

Regarding the viewing conditions, five essential elements are taken into consideration. Firstly, Majuddin et al. (2021) suggest that the number of viewings of the selected excerpt may determine L2 acquisition; that is, watching the audiovisual materials more than once enhances meaning understanding and leads to higher comprehension, especially if supported with captions, as demonstrated in Winke et al. (2010, cited in Majuddin et al., 2021) and Chang & Read (2006, cited in Majuddin et al., 2021)'s studies. The reason behind this, according to Hatch (1983, cited in Chang & Read, 2006, p.378), is that repetition "provides more processing time and clarifies the relationship of syntactic forms"; however, repetition can involve either replaying the whole input or repeating certain excerpts to simplify it. Nonetheless, Chang and Read (2006) emphasise the importance of identifying the learners' proficiency level in order to determine the number of

viewings required for input comprehension; that is, the higher the listening proficiency, the fewer the repetitions needed.

Secondly, whether captions are used during the first or the second viewing is considered an influential factor for input recognition (Montero Perez, 2022). According to Winke et al. (2010), a first viewing with on-screen text eases learners' anxiety and activates innate strategies for language selection and processing, leading students with captions first to better recognize L2 vocabulary than students with captions second. In fact, participants from Winke et al. (2010)'s study stated that the use of captions in the first viewing foregrounded the written input and allowed them to identify the target language in the video; hence, in the following viewing, they already knew what to pay attention to – attention has a key role in L2 learning, seeing that learners need to consciously target the input to successfully acquire it (Schmidt & Frota, 1986, cited in Winke et al., 2010). To put it another way, when captions occur first, students' attention is drawn to new language, allowing to gather information in the second viewing; conversely, when captions occur second, students identify unknown language in the first viewing and verify it with the aid of captions (Gass, 1988, cited in Winke et al., 2010). In order to make a decision on the order of captions, however, the orthographic distance between L1 and L2 should be taken into account; that is, if the L1 and L2's orthography are close (e.g., Spanish and English), captions should be used in the first viewing, whereas if they are distant, captions should be used second (Gass, 1997).

The third viewing condition identified is the pace of audiovisual materials. According to Mayer et al. (2014), fast-paced excerpts are more difficult to comprehend than slow-paced videos regardless of the use of captions, seeing that L2 learners need time to process the input. On the case that on-screen text is used, fast-paced videos remain challenging for learners because of the captions' speed, that is, students need time to carefully read captions and comprehend them; in fact, Mayer et al. (2014) point out that the lack of speed-reading skills inevitably leads to missing input. Slow-paced videos, hence, become useful especially for non-native speakers, as they are given time to carefully read and process the written support for better language comprehension.

Finally, the two remaining conditions inevitably overlap. On the one hand, narrow viewing consists of watching several episodes of the same series or program, hence allowing a frequent exposition to similar, thematically-related words (Majuddin et al., 2021). This technique is already beneficial for L2 students, seeing that episodes provide context for the following episodes, which fosters input comprehension; however, the use of captions becomes an extra support for students, especially in the first or in the most challenging episodes (Rodgers & Webb, 2017). On the other hand, the viewing length can also determine the effectiveness of captioned audiovisual materials for language acquisition. As explained by Montero Perez (2022), a distinction can be made between the

length of the excerpt and the number of sessions involved: whereas single sessions can be dedicated to watching either short clips between 9 and 20 minutes or full episodes between 20 and 60 minutes, multiple full episodes can be watched throughout several sessions, as in narrow viewing, ranging from 7 up to 24 sittings. In fact, studies have proven that viewing length can determine the acquisition of L2 vocabulary when comparing pre-tests and post-tests; however, notwithstanding these gains, Montero Perez (2022, p.175) points out that other factors should be taken into account, including "word-related and learner-related variables".

2.4.2 Techniques applied to captions

As regards the techniques applied to captions, there are several essential methods that should be taken into consideration. To begin with, keyword captioning is a means to make certain words stand out from the rest of the on-screen text (Majuddin et al., 2021); that is, rather than presenting the full text in the excerpt, only representative words of the input are included in order to draw attention to specific content (Montero Perez, 2022). Guillory (2013) advocates the use of keyword captions by demonstrating in his study their effectiveness for L2 listening comprehension. The researcher displayed centralized three-word captions – according to Schlesinger (1968, cited in Guillory, 2013, p.94), "readers read in chunks of 2.5 to 3 words" – that participants could easily identify and process; as hypothesized, there was no significant difference between the full-captions group and the keyword-captions group, thus successfully proving not only the efficacy and benefits of keywords, but also the dispensability of full on-screen text. However, other researchers such as Montero Perez (2022, p.170) differ from Guillory (2013)'s claims by stating that some learners might find keywords "distracting", as well as considering results of studies on keyword captions effectiveness "inconclusive".

Similarly, highlighting words is a useful technique to make the targeted input more salient with the aim of improving learners' immediate new word form and meaning recall abilities (Cintrón-Valentín & García-Amaya, 2021; Montero Perez, 2022). Researchers Lee and Révész (2020) noticed the effectiveness of highlighted captions for introducing new, targeted lexical units to L2 students, leading to inevitable learning gains; in fact, many studies have demonstrated the benefits of highlighting target words in captions, especially for adult intermediate-level L2 learners (Montero Perez, 2022). Nonetheless, other researchers consider that the effectiveness of highlighted words for L2 learning cannot be proven, as it rather "detracts [students'] attention from other elements" (Majuddin et al., 2021, p.989); in other words, salient words get to occupy the foreground at the expense of the unenhanced input, which becomes impossible to recall for students (Choi, 2018, cited in Majuddin et al., 2021). Apart from the effectiveness of highlighted words,

Cintrón-Valentín and García-Amaya (2021)'s study puts into question the role of other typographic enhancement techniques, bolding and underlining, for L2 learning of targeted input. Their study, among many others, has proven the positive outcome of on-screen text enhancement; however, they have all concluded that, despite the benefits of bold and underlined captions for vocabulary acquisition, they were no more useful than captions alone, that is, "the presence of typographic enhancement did not lead to significantly better recall compared to viewing with normal captions" (Majuddin et al., 2021, p.1001).

Another means to foster learners' ability to connect form and meaning is the use of glossed captions, which are annotations added to the on-screen text not only to draw the learners' attention to target lexical units, but also to clarify their meaning; that is, glosses provide immediate access to the input's meaning and translation, whether it be on the margin of the screen or by directly clicking on the target word (Boers, 2022, & Montero Perez, 2022). Many researchers including Cheng and Good (2009, cited in Boers, 2022), Ko (2012, cited in Boers, 2022) and Boers (2022) consider this technique to be beneficial for vocabulary acquisition and text comprehension, and it has even been found to be more effective than the previously mentioned methods, keyword captioning and typographic enhancement (Montero Perez, 2022). However, the effectiveness of glossed captions is possible as long as the input is comprehensible; in other words, learners will not be able to benefit from them unless the input is understood (Boers, 2022). Besides, Abraham (2007) points out the importance of learners not only consulting the glosses, but also not spending most of the time reading them, as some studies have found out that the lack of input comprehension is due to the fact that participants either completely ignore the glossed captions or focus so much on them that forget about the rest of the input.

The last technique to mention is bilingual subtitles, which consists on presenting the on-screen text both in the L2 and L1 simultaneously (García, 2017) to foster not only L2 vocabulary acquisition, but also form-meaning recall (Li, 2016, cited in Wang & Pellicer-Sánchez, 2022). Some studies such as García's (2017) point out the accessibility to the grammatically-correct written form of L1 and L2 input enabled by dual-captioned media, allowing learners to improve their vocabulary; however, other researchers claim that the L1 subtitles may distract learners and keep them from processing the L2 new lexical units (Peters, 2019, cited in Wang & Pellicer-Sánchez, 2022). As Montero Perez (2022) suggests, participants may even ignore one of the languages and rely on one source only, especially in the L1. Likewise, notwithstanding the benefits of bilingual subtitles for meaning recognition and recall, Wang and Pellicer-Sánchez (2022) state that participants focusing on the L1 subtitles only leads to poor performance on the post-tests as regards form recognition of L2 target input. Hence, the effectiveness of dual-captioned media for L2 acquisition remains

inconclusive, as the few studies existing have yielded mixed results (Montero Perez, 2022; García, 2017; Wang & Pellicer-Sánchez, 2022).

2.5 Studies on captioned and non-captioned conditions

Many researchers have attempted to prove the benefits of captions for L2 acquisition by conducting different studies on participants divided between captioned-condition (henceforth, CC) and non-captioned-condition (henceforth, NCC) groups. Some of the linguistic features that have been examined include speaking skills, listening comprehension and vocabulary acquisition, among others, as is explained in what follows.

As regards speaking performance, Kim (2020, p.151)'s experiment comprised of 67 university students of English from southern Korea between the ages of 19 and 22. All participants took a test to determine their L2 proficiency and were subsequently divided between low-level and high-level proficiency groups. Having divided each group into CC and NCC, all participants were exposed to 10 episodes of the American sitcom *Friends* throughout 10 sessions, and took a mid-term and a final speaking tests. Whereas the low-level proficiency CC participants performed slightly better than the NNC participants as regards fluency, accuracy and complexity, the high-level proficiency CC participants significantly outperformed the NNC participants, especially in terms of fluency and accuracy. Hence, the researcher concluded that, despite the effectiveness of on-screen text for both low- and high-level L2 learners, captions are especially beneficial for proficient speakers to produce "spoken language more fluently and more accurately".

Regarding listening comprehension and vocabulary acquisition, Hsu et al. (2013) evaluated a group of 81 11-year-old students of English at an elementary school in Taiwan for a month. Three classes of beginner proficiency level became a different group each, two of which were CC and NCC, and they had to use their personal digital devices to watch selected videos related to the class lesson. The participants were tracked of their L2 learning progress on a weekly basis with a test immediately after each viewing and, despite both groups made progress, the CC participants outperformed the NCC group, especially after the third week. In fact, there was no difference between the listening comprehension between them, but it was the vocabulary acquisition that was significantly greater for the CC group.

Likewise, Jelani and Boers (2018) conducted their study on a group of 81 16-year-old high school students from Malaysia who were divided into three groups, two of which were CC and NCC, to test their vocabulary acquisition. The participants, all intermediate-level speakers of English, watched a 10-minute TED talk thematically related to one of the topics of the course textbook, and took written and aural word-recognition and word-meaning post-tests immediately

after the viewing. The researchers determined that the results of the word-recognition tests were inconclusive, as there was no significant difference between the CC and NCC groups; however, the first outperformed the latter in the written word-meaning test, but not in the aural. Hence, Jelani and Boers concluded that, despite the benefits of captions for vocabulary acquisition, test modality has to be taken into account.

By contrast, Bisson et al. (2014) argue in their experiment that neither CC nor NCC are effective for L2 vocabulary acquisition. The study originally comprised 64 participants that took a questionnaire about their linguistic backgrounds, but, having excluded unqualified participants, in the end only 54 of them were part of the experiment. The task consisted on watching a 25-minute excerpt of the animated film *SpongeBob Square Pants* and, immediately after, taking an aural vocabulary test of 156 one-word items. The researchers observed that there was no difference between the results of the CC and NCC groups and, hence, they concluded that there was no evidence of the effectiveness of captions for L2 vocabulary acquisition. However, it is questioned whether the outcome of their experiment might be due to some limitations such as the short exposure to the L2 or the lack of L2 previous knowledge (Jelani & Boers, 2018) – the participants were English native speakers that had no knowledge of Dutch.

2.6 Limitations of written support in audiovisual materials

When it comes to written support, some criticism has arisen concerning the use of text in clips. Paivio (1986) states in his *Redundancy Principle* that the simultaneous use of visual, hearing and written support does not facilitate learning but rather interferes with it (cited in Montero Perez et al., 2013), seeing that subtitles and captions may draw the students' attention from the audiovisual material (Kaiser, 2011, p.239); as Frumuselu et al. (2015, p.109) put it, "one might believe learners stop listening to the original soundtrack while reading their native language".

Despite the large number of studies on language acquisition through audiovisual materials, Montero Perez (2022) mentions the fact that the wide variety of scope and methodology between studies sets back comparisons that allow to determine vocabulary learning, thus leading to learning gains being unclear. Besides, Van Lommel et al. (2006, cited in Montero Perez, 2022)'s investigation determined that captions in audiovisual materials are not as effective for grammar acquisition, taking into account that grammatical rules require explicit instruction prior to watching the excerpt; hence, having no formal explanation, for incidental learning of grammatical structures to happen, the researchers suggest that learners should be frequently exposed to the targeted grammar – Pattemore and Muñoz (2020)'s experiment applies this suggestion by aiming for learners to acquire grammar after watching 10 captioned episodes of a TV series.

Nonetheless, there are still areas to study and research. Barón and Celaya (2022) state that such studies are limited timewise, that is, they have to be done over a long period of time. In order to analyse whether students have retained the input, they should be tested after a certain period of time rather than right after viewing the audiovisual materials. In addition, Montero Perez et al. (2013) mention that most studies are conducted on low-level L2 learners and, hence, research on the effectiveness of written support for advanced or proficient learners is still yet to be explored. Likewise, vocabulary acquisition as the main focus of research, studies on L2 grammar and syntax learning through videos with on-screen text are scarce (Montero Perez, 2022).

3 THE STUDY

3.1 Context

The setting in which the experiment was conducted is a language school located in Girona, Spain, in which the participants attend weekly English classes; specifically, 3 hours on Friday evenings from September 2022 to June 2023 with a native British teacher. The group has a C2 level of English and is preparing to take the Cambridge Proficiency exam in order to obtain the Cambridge Certificate of Proficiency in English (CPE), which is considered the highest qualification in the English language as holders of this certificate have native-like fluency of English.

3.2 Participants

The study involved a group of proficient-level English students (n=6) between the ages of 16 and 23, all of whom were female (LA, CR, NE, CL, and MI) except for one male (AR), and they are currently studying either at high-school or university. The participants are Catalan-Spanish native speakers who fluently speak at least another foreign language apart from English, such as French, Italian or Chinese. Whereas some of them only are exposed to English during the 3-hour class at the language school, others are involved with the language for over 7 hours per week; besides, despite 5 of them having visited English-speaking countries, only CR has actually lived in one, the USA. Nonetheless, the only participant who has not been abroad has had a live-in English au-pair since she was a child.

As regards their watching habits, they all claim to watch films and series between once and thrice a week, always in their original language, whether it be English, Spanish, Catalan, or even French. Out of all the participants, only four of them use subtitles or captions in native-English audiovisual materials; however, they state to always use English captions rather than an on-screen translation. Besides, whereas three of them would rather watch audiovisual materials in American English and another participant prefers British English, the two remaining claim that they do not a

preference for one over the other. Finally, it is important to note that none of them have ever watched *Red Rose*, the Netflix series of which the 5-minute excerpt is extracted; in fact, they had never heard of it before.

3.3 Instruments

In order to proceed with the experiment, the participants had to fill in a questionnaire (see Appendix 1) created on Google Forms which included 22 closed-ended and open-ended questions to determine their sociolinguistic backgrounds. The questions aimed to collect information not only on the participants personal information, spoken languages and proficiency, and their English learning experiences, but also their television watching habits, as well as their use of subtitles and captions. Besides, they had to fill in a consent form (see Appendix 2) to confirm they agree to participate in the study and that their information will remain anonymous.

Moreover, the participants had to watch a 5-minute video excerpt of the Netflix horror series *Red Rose*, to be more concrete, its first episode, which presents the lives and conversations of a group of 17-year-old friends from Bolton, UK; hence, the audiovisual material involves numerous British slang terms and idioms that can be targeted. The excerpt, extracted from the BBC Three website and edited with Clipchamp, was offered to the participants either with CC or NNC, depending on their group, both containing the original English audio.

Finally, to analyse the effectiveness of captions for L2 acquisition, a test (see Appendix 3; for the test key, see Appendix 4) was created and distributed to the participants both before and after the watching of the excerpt. It consisted of 15 questions on British slang terms and idioms, 11 of which were target vocabulary extracted from the series, whereas the other 4 were mere distractors, that is, vocabulary that does not appear in the video. Besides, the test included 8 closed-ended questions in which the participants had to choose the correct meaning of the targeted expression, as well as 7 open-ended questions in which the meaning, synonym or translation of the targeted expression had to be provided by the participants; in fact, the main purpose underlying the decision of including a combination of multiple-choice questions and open questions is avoiding participants to guess the answers.

3.4 Procedure

The study took place at the language school (see 3.1) during class time throughout two different sessions. The first, held on the 14th of April, consisted on the participants filling in the sociolinguistic questionnaire and signing the consent form. On the one hand, the questionnaire consisted of questions related to three main categories: demographic information (e.g., age,

languages, and education), out-of-class exposure to L2 (e.g., number of hours, countries visited, and living abroad), and watching habits (e.g., number of hours, languages used, and use of subtitles/captions). To make it easier to collect all the information, the questionnaire was distributed during the first 5-15 minutes of the class. Having collected all the answers, they were analysed, not only to know about the participants' sociolinguistic background, but also to make sure none of them had previously watched *Red Rose*, the series from which the 5-minute excerpt was extracted. On the other hand, the consent form was distributed on paper so participants could sign it and collected during the first minutes of class. By signing it, they agreed to participate in the study and understood that their information would remain anonymous.

During the second session, held a week later, on the 21st of April, the link for the pre-test, created on Google Forms for easier collection of the answers, was shared on the WhatsApp group so that participants could access it immediately. They were explained that they had to answer the test within 10 minutes and, most importantly, not to go to the next page in the form until everyone had finished; besides, they were told not to look for the answers they did not know nor try to guess them – I was present when they were completing the test to ensure that nobody cheated. After the 10 minutes they were given to complete it, they were allowed to go to the next page in order to watch the 5-minute video of the Netflix series *Red Rose*. The participants were randomly assigned to two different groups, one with CC (NE, AR, MI) and the other with NCC (CL, LA, and CR), with an online random generator; hence, they were instructed to watch only the link of the video corresponding to their group and, once again, to not go to the following page until everyone had finished. Finally, immediately after the viewing, they were given 10 more minutes to take the post-test (see 8.3 and 8.4), which included the exact same questions as the pre-test, and were reminded of the same instructions of the pre-test; that is, they were told not to look on the Internet nor try to guess the answers. However, they were encouraged to submit the test as soon as they finished, seeing that that was the last stage. The whole procedure took between 30 and 40 minutes in total: between 5 and 15 in the first session, and 25 in the second.

4 RESULTS

The research question of the present study focused on which of the two conditions, captioned and non-captioned, is more effective for the acquisition of British slang terms and idioms; in order to answer it, the results from both the pre-test and the post-test were analysed with raw numbers, averages and percentages.

The independent variables in the study are the CC and NCC, whereas the dependent variables are the answers in the tests, which were divided between "correct", "incorrect" or "do not

know" (henceforth, DK). To determine whether an answer could be considered valid or not in open-ended questions, they were carefully considered according to their degree of accuracy.

4.1 Captioned-condition group

The results of the CC group (n=3), formed by CR, AR, and MI, demonstrate that their overall performance has significantly improved from the pre-test to the post-test. On the one hand, the pre-test scores show that participants in the CC group either did not know many of the answers or decided not to take a guess, as a 75% of slang terms and 66.66% of idioms were not answered (see Table 2) – in fact, the average was highly affected by MI's answers, as the participant did not answer 14 out of the 15 questions, and the remaining one was incorrect. However, of those questions that were answered, there was only a 2.77% difference between correct (see Table 1) and incorrect (see Table 2) slang terms; that is, whereas a 13.88% were wrong, only a 11.11% were right. As regards idioms, there was a difference in performance, seeing that a 22.22% of idioms were correct and only a 11.11% were incorrect. Hence, the overall score of correct answers in the pre-test was a 16.67%.

On the other hand, the post-test scores show a great improvement in the performance of participants in the CC group, specially taking into account that the percentage of unknown answers was reduced from 75% to 41.66% for slang terms, and from 66.66% to 33.33% for idioms. Of those questions that were answered, most were correct, as the percentage of accurate responses increased from 11.11% to 36.11% for slang terms, and from 22.22% to 55.55% for idioms. However, the number of incorrect slang terms also rose from 13.88% to 22.22%, despite the percentage of erroneous idioms remaining the same (11.11%). Nonetheless, the overall performance of participants after watching the captioned 5-minute video grew from 16.67% to 45.83%.

Table 1. Number and percentage of correct answers of each CC-group participant.

Group	PRE-TEST				POST-TEST			
	Slang terms		Idioms		Slang terms		Idioms	
	Out of 12	%	Out of 3	%	Out of 12	%	Out of 3	%
CR	3	25	2	66.66	5	41.66	3	100
MI	0	0	0	0	4	33.33	1	33.33
AR	1	8.33%	0	0	4	33.33	1	33.33
Average	1.33	11.11	0.66	22.22	4.33	36.11	1.66	55.55
Final % average	16.67				45.83			

Table 2. Number of incorrect and "don't know" answers of each CC-group participant.

	PRE-TEST				POST-TEST			
	Slang terms (out of 12)		Idioms (out of 3)		Slang terms (out of 12)		Idioms (out of 3)	
Group	Incorrect	DK	Incorrect	DK	Incorrect	DK	Incorrect	DK
CR	4	5	0	1	6	1	0	0
MI	0	12	1	2	0	8	1	1
AR	1	10	0	3	2	6	0	2
Average	1.66	9	0.33	2	2.66	5	0.33	1
Average in %	13.88	75	11.11	66.66	22.22	41.66	11.11	33.33

4.2 Non-captioned-condition group

The results of the NCC group (n=3), formed by LA, NE, and CL, demonstrate that the participants have slightly improved their performance from the pre-test to the post-test. In the test previous to the video excerpt, the participants answered correctly a 19.44% of the questions: a 16.66% of slang terms and a 22.22% of idioms were accurate (see Table 3). Of the remaining questions, a 36.11% of slang terms and a 44.44% of idioms were incorrectly answered (see Table 4), whereas a 47.22% of slang terms and a 33.33% of idioms were not answered at all (see Table 4).

As regards the post-test, the scores show a slight improvement in the overall percentage of correct answers, from a 19.44% in the pre-test to a 38.89% in the post-test; both the slang terms and idioms had more accurate answers, rising from 16.66% to 33.33% for the former, and from 22.22% to 44.44% for the latter. Hence, the average percentage of incorrect and unknown answers slightly decreased. Whereas the incorrect slang terms and idioms were reduced from 36.11% to 33.33% and from 44.44% to 33.33%, respectively, the percentage of unknown L2 vocabulary decreased from 47.22% to 33.33% for slang terms and from 33.33% to 22.22% for idioms. Therefore, participants scored the same percentage of slang terms in each of the categories during the post-test (i.e., 33.33%), whereas idioms had more correct answers, followed by incorrect answers and finally unknown answers.

Table 3. Number and percentage of correct answers of each NCC-group participant.

	PRE-TEST				POST-TEST			
	Slang terms		Idioms		Slang terms		Idioms	
Group	Out of 12	%	Out of 3	%	Out of 12	%	Out of 3	%
LA	1	8.33	1	33.33	4	33.33	1	33.33

NE	1	8.33	0	0	1	8.33	1	33.33
CL	4	33.33	1	33.33	7	58.33	2	66.66
Average	2	16.66	0.66	22.22	4	33.33	1.33	44.44
Final % average	19.44				38.89			

Table 4. Number of incorrect and "don't know" answers of each NCC-group participant.

	PRE-TEST				POST-TEST			
	Slang terms (out of 12)		**Idioms (out of 3)**		**Slang terms (out of 12)**		**Idioms (out of 3)**	
Group	**Incorrect**	**DK**	**Incorrect**	**DK**	**Incorrect**	**DK**	**Incorrect**	**DK**
LA	3	8	0	2	2	6	0	2
NE	6	5	3	0	7	4	2	0
CL	4	4	1	1	3	2	1	0
Average	4.33	5.66	1.33	1	4	4	1	0.66
Average in %	36.11	47.22	44.44	33.33	33.33	33.33	33.33	22.22

4.3 Comparison between CC and NCC groups

Having analysed both groups, the significant difference in performance between the pre-test and post-test is most noticeable (see Figure 1); that is, despite the fact that both groups obtained similar results in the pre-test (16.67% in the CC group and 19.44% in the NCC group), the gap has widened in the post-test results, as the CC group obtained a 45.83% of correct answers, improving their performance by 29.16%, whereas the NCC group got a 38.89% of accurate answers, which is only a 19.45% increase. Hence, there is almost a 10% difference between both groups.

As regards unknown and incorrect answers, almost all of them have decreased in percentage (see Figure 2). On the one hand, there was a 30.55% gap between the unknown answers of the CC and NCC groups in the pre-test, as they did not answer a 70.83% and 40.28% of the questions, respectively. However, in the post-test the percentages of unknown answers were reduced to 37.5% in the CC group and to 27.78% in the NCC group; hence, at the end there was only a 9.72% gap between them. On the other hand, whereas the NCC group has reduced their incorrect answers by almost a 7%, from a 40.28% in the pre-test to a 33.33% in the post-test, the CC group has actually gotten a higher percentage of erroneous answers, increasing the gap by a 4.17%, from 12.5% in the pre-test to 16.67% in the post-test. Nonetheless, the percentages in the CC group have remained much lower than in the NCC, as there is a gap of 27.78% in the pre-test and of 16.66% in the post-

test between both groups.

Figure 1. Percentage of correct answers of CC and NCC groups both in the pre-test and post-test.
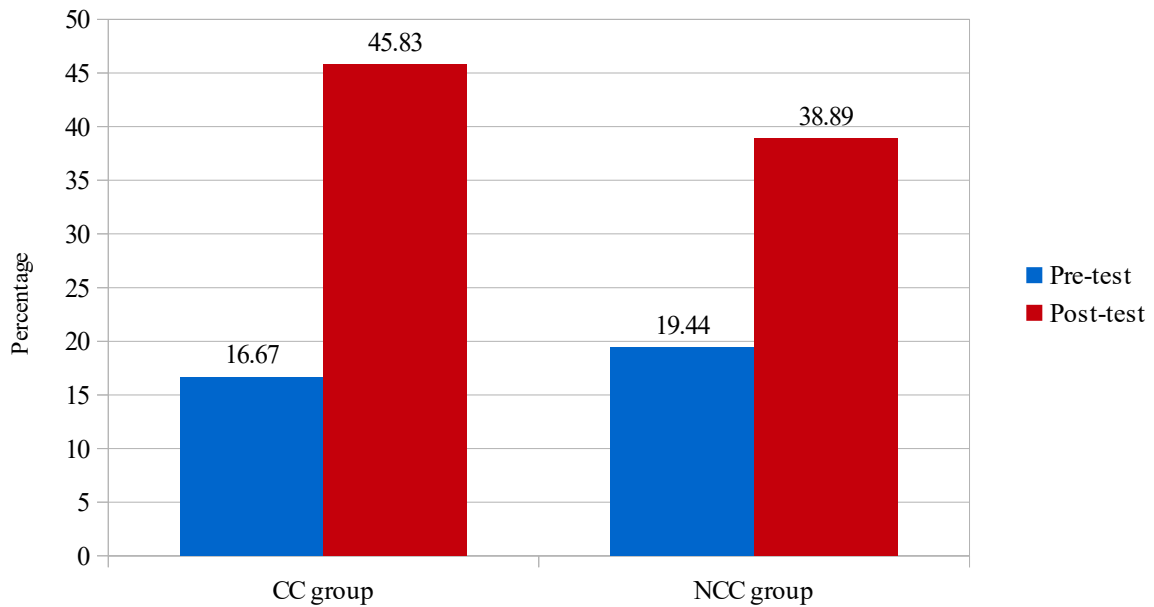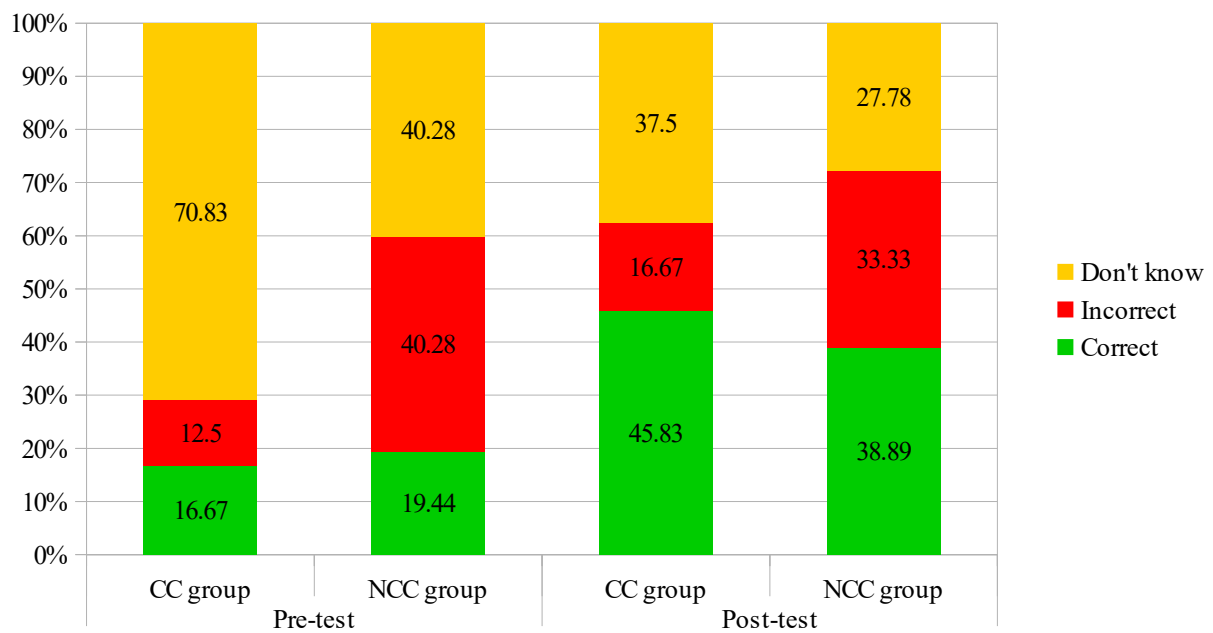

Figure 2. Correct, incorrect and unknown answers of participants in percentages. The data is divided between pre-test and post-test answers for both CC and NCC groups.

Furthermore, it is worth mentioning that, independently of the group condition, there were some specific slang terms that neither of the groups were able to answer correctly (see Table 5), to be more concrete, *tenner* and *youse*. Both of them were unanswered by everyone in the pre-test, as nobody knew what they meant; in the post-test, two people answered *tenner* and one of them also answered *youse*, however, the three were incorrect. To put it another way, regardless of the captioned or non-captioned conditions, none of the participants were able to understand the meaning of these terms. Likewise, no one was able to explain the meaning of *summat* in the pre-test; however, although there were four incorrect and unknown answers, that is, two of each, two participants from the CC group were able to answer correctly. Hence, the comprehension of this slang term was aided with captions.

Finally, despite the CC group outperforming the NCC group in the post-test, there are some questions in which the latter has gotten more positive results. For instance, the words *tight*, in the context *Don't be tight*, and *scrote,* in the context *Rather stand with a scrote than a slaggy scrubber*, have been correctly answered by two NCC-group participants but only one from the CC group; similarly, the expression *Do one* has only been scored by one of the students, who belonged to the NCC group. In fact, the NCC group has even gotten better scores in distractors such as *Take the Mickey out of someone*, which means to "make fun of someone"; that is, despite neither appearing in the video excerpt nor having on-screen written support, two out of three participants of the NCC group have correctly answered the question, whereas only one of the CC group has gotten it right.

Table 5. Difficult British slang terms for participants.

Slang term	PRE-TEST (out of 6)		POST-TEST (out of 6)	
	Incorrect	DK	Incorrect	DK
Tenner	0	6	2	4
Summat	1	5	2	2
Youse	0	6	1	5

5 DISCUSSION

The aim of the current study was to determine which condition, CC or NNC, was more effective for the acquisition of L2 vocabulary, specifically British slang terms and idioms. As shown in the results (see section 4), and in line with Barón and Celaya (2022)'s findings, the use of audiovisual materials for the acquisition of L2 vocabulary is effective regardless of whether on-screen text is used or not; in other words, the comprehensibility and acquisition of new L2 vocabulary has been

fostered for both CC and NCC groups after being exposed to a multimodal form of input, that is, aural and visual, as Collins & Muñoz (2016, cited in Muñoz, 2022) mention in their study. As explained in section 4, both groups have significantly increased their performance levels, seeing that the CC group scored almost half of the questions right (i.e., 45.83%) whereas the NCC group almost doubled theirs (i.e., 38.89%). Hence, evidence illustrates that contextualization of L2 vocabulary through audiovisual materials allows a better comprehension of the input and, consequently, they become useful learning resources, as suggested in several studies (Montero Perez, 2022, Elgort et al., 2018, & Lin and Siyanova-Chanturia, 2015, cited in Montero Perez, 2022).

Nonetheless, post-test results prove that the addition of written support during the 5-minute watching encourages the development of participants' reading and listening skills, as well as the internalization of L2 vocabulary, as stated by researchers such as Frumuselu et al. (2014) and Montero Perez et al. (2013). To put it another way, participants in the CC group, who had the input both in English audio and English transcription, outperformed participants in the NCC group, who only had the aural input, seeing that they scored a higher percentage of correct answers in the post-test, despite the difference between both groups being minimal (i.e., 9.71%; see Figure 2). In fact, the slight difference in post-test results might be determined by the fact that captions can become a mere textual support for proficient L2 learners, as suggested by Frumuselu et al. (2015), seeing that, rather than encouraging L2 acquisition, they might simply allow participants to visualize what they hear. Hence, taking this into account, conducting the study on low or intermediate-level students might have resulted in a higher percentage of CC-group correct answers, since researchers such as Pujadas & Muñoz (2019) and Montero Perez et al. (2013) mention captions allow beginner learners not only to reinforce their listening comprehension, but also to succeed in language acquisition. Nevertheless, while intermediate-level participants would have beet suitable for the study, Vanderplank (1988) suggests that beginner students can benefit more from subtitles rather than captions.

Although results have proven the benefits and effectiveness of captions for L2 acquisition, there are researchers who, on the contrary, claim that on-screen text might interfere with learning. Montero Perez et al. (2013) mention in their study that, according to Paivio (1986)'s *Redundancy Principle*, captions do not facilitate L2 learning, but they rather distract participants by drawing their attention from the visual input; that is, learners might focus on reading the text rather than acquiring new vocabulary. Likewise, Kaiser (2011) and Frumuselu et al. (2015)'s results, among other researchers, state that learners exposed to multimodal forms of input will unconsciously prioritize the audio over the written support, that is, either subtitles or captions. Nonetheless, in

order to study whether participants spent more time reading the text or listening to the audio and, subsequently, compare the results with the scored answers in the vocabulary tests, an eye-tracking study should have been conducted; that is, many researchers such as Bisson et al. (2014) conduct experiments consisting of recording the movement of the participants' eyes in order to determine whether their percentage of correct answers in the post-test is due to the contextualization of the input through audiovisual materials or due to the time spent reading the captions.

As regards incorrect answers, Figure 2 (see section 4.3) shows that, contrary to the NCC group, the CC group increased the average percentage of erroneous answers by 4.17% in the post-test rather than decreasing it; however, Table 2 (see section 4.1) proves that only 3 more slang terms had been incorrectly guessed in comparison to the pre-test. In other words, the fact that the study was constituted by six participants only inevitably leads to the whole group's average decreasing when one of the participants gets a wrong answer, that is, the percentage of incorrect answers will be highly affected even by minimal errors. Hence, taking into account the inverse proportionality of these variables, most researchers decide to test a large number of participants in order to obtain significant results, as can be seen in Kim (2020), Hsu et al. (2013), Jelani and Boers (2018), and Bisson et al. (2014)'s studies.

The present study's results, however, are not in line with Jelani and Boers (2018)'s findings. Aiming to determine the effectiveness of captions for L2 vocabulary acquisition, the researchers tested 81 16-year-old participants who were, at the time, attending classes at a Malaysian high school and all of them were intermediate-level speakers of English. The students, divided into a CC and a NCC group, had to watch twice a 10-minute TED talk thematically related to the coursebook and were given immediately after a multiple-choice comprehension test on the video. Then, a post-test containing 15 target words and 13 distractors, and divided into three different parts, was distributed; the tasks tested the participants' abilities to recognize words, both written and orally, the orthographic accuracy of target words, and the meaning comprehension of both target vocabulary and distractors. Having analysed the results, the researchers deemed the results inconclusive, seeing that there was no significant difference between both groups; however, not in the aural, but in the written word-meaning test, the CC group actually outperformed the NCC group. This emphasises the importance of test modality in linguistic studies, as expressed by Majuddin et al. (2021), since findings may vary among studies due to the modality of test chosen; that is, if the present study had involved an aural rather than a written test, or perhaps the aim was, for instance, to test meaning understanding, orthographic accuracy, listening comprehension or speaking performance, the results could have been completely different, seeing that proficient learners of English have more expertise in some areas of the language. To put it another way, despite their fluency in the L2, some students

might outperform others in grammar and vocabulary, whereas others might stand out in speaking and listening tasks; hence, other areas of language could be explored in future studies. Likewise, although the age range was only slightly different from that of the present study, the number of participants, as well as their proficiency level of English and number of viewings, are determining factors for the outcome of the post-test results. Majuddin et al. (2021) suggest in their study that watching the selected video more than once, especially if on-screen text is included, encourages language acquisition and leads to higher comprehension; besides, it has to be taken into account not only that the use of captions can benefit more post-intermediate-level learners of English, but also that their fluency in the L2 will allow them to comprehend and acquire different target input. Nonetheless, Montero Perez et al. (2013) mention in their study that there is a lack of research on the effectiveness of written support for proficient-level students as most studies involve beginner and intermediate learners of the L2 and, hence, more proficient participants should be involved.

Notwithstanding the fact that different factors may lead to diverse outcomes, the present study remains in line with the findings of other studies despite not sharing most characteristics, such as Frumuselu et al. (2015)'s. The mentioned researchers' study aimed to determine the effectiveness of both subtitles and captions for the acquisition of informal and colloquial language by exposing learners to several episodes of the American sitcom *Friends*. However, despite them reaching the conclusion that captions are more effective for the acquisition of L2 vocabulary, as has been proved by the present study as well, the time of exposure in Frumuselu et al.'s experiment was much longer, as it consisted of 25-minute sessions held throughout seven consecutive weeks, to be more concrete, two sessions per week. This technique of watching several episodes of the same series, called narrow viewing, is considered by Rodgers & Webb (2017) to be beneficial for students not only for the context provided by the episodes, but also for the frequent exposure to thematically-related vocabulary. That is, being exposed to numerous episodes of a series rather than to a 5-minute excerpt, the participants are able to comprehend the input better, leading to an improvement in vocabulary acquisition. In fact, Barón and Celaya (2022) mention in their study that time can be a limitation in experimental research, seeing that, in order to determine whether the participants have retained the input or not, such studies should be done over a long period of time. Hence, the findings of the present study might have shown a wider gap between CC and NCC groups' results having exposed the students to the whole episode of which the video excerpt was selected, or even to several episodes of the series; yet a 5-minute video was selected taking into account Bruti (2016)'s emphasis on the selection of short clips rather than whole episodes in order to allow learners to focus on the language and, hence, analyse and understand it better (Kaiser, 2011). Besides, an important feature to bear in mind is the English variety of the video, seeing that the

present study involved British slang terms and idioms, whereas Frumuselu et al.'s consisted of American informal and colloquial language. That is, taking into account that most English-native multimedia is generated in the USA, most students are used to American English and, hence, British English, specifically the Lancashire dialect, might be more difficult to understand. Thus, having chosen an American series rather than a British one, results might have differed.

6 CONCLUSIONS

The present study aimed at examining whether the use of captions or no captions in audiovisual materials had an effect on the acquisition of L2 vocabulary, more specifically, British slang terms and idioms. Researchers, such as Montero Perez et al. (2013) and Barón & Celaya (2022), emphasise the effectiveness of multimedia as a useful language acquisition resource, regardless of the addition of written support; the present study's results (see section 4) seem to be in line with their findings, as both the CC group and the NCC group showed a slight improvement in the comprehension of new vocabulary in the post-test. Hence, it can be stated that the contextualization of the input through audiovisual materials fostered language learning regardless of the addition of on-screen text. In fact, as mentioned in previous studies, teachers should not only introduce such resources into the classroom, but also encourage L2 students to use multimedia as an incidental learning tool both inside and outside the academic context (Frumuselu et al., 2015), seeing that technology can result in the overlapping between formal and informal learning as traditional educational methodologies and learning drift apart. Thus, audiovisual materials can be considered essential in the acquisition of second languages.

Nonetheless, the implementation of written support can further enhance the development of language skills, such as vocabulary comprehension and acquisition. Although it can be divided between subtitles and captions, only the latter was involved in the present study in order to align with the proficient level of English of the participants; to put it another way, seeing that students with a high level of L2 proficiency can benefit more from captions rather than subtitles, as suggested by Vanderplank (1988), the present study aimed to analyse the effectiveness of on-screen text transcriptions for L2 acquisition. In fact, the post-test results certainly demonstrated that the CC group significantly improved their performance in comparison to the pre-test, scoring almost half of the vocabulary after having watched the video excerpt and, hence, widening the gap between NCC and CC groups by almost a ten percent; in other words, the findings proved the effectiveness that the addition of written support has on the acquisition of new vocabulary as the CC group outperformed the NCC group in the post-test despite the former obtaining poor results in the pre-test.

The study, however, has some limitations that should be addressed. Despite most researchers highlight the importance of conducting experiments on a great number of participants, only six were involved in the present study, seeing that on the one hand, they were easily available, and, on the other hand, finding participants with similar sociolinguistic backgrounds would have been a difficult task; hence, not only to determine the effectiveness of captions for the acquisition of L2 vocabulary but also to increase the reliability of the present study, future experiments should be comprised of a larger number of participants, at least between fifty or eighty. Besides, the video watching lasted for 5 minutes only, being able to conduct both the pre-test and post-test in the same session as the viewing; though, in order to examine thoroughly the effects of captions on the acquisition of L2 vocabulary, future studies should consist of several episodes of the same series and should be conducted over a long period of time. To put it another way, narrow viewing should be considered in further studies and students should be tested after a while in order to examine whether they retained or not the input. Finally, an eye-tracking study could be carried out in order to analyse whether the improvement of post-test results in the CC group were a consequence of partially or completely reading the on-screen text and, at the same time, which words were more often read. However, the possibility of holding such study was immediately discarded seeing that training on how to use the eye-tracker machine is required and time was limited; hence, it was not possible nor a priority to develop this ability.

All things considered, the findings of the present study should be considered by teachers in the decision of including new approaches to language acquisition in the class curriculum, in particular of colloquial and informal vocabulary, which is rarely present in the academic context. In fact, the post-test results of both CC and NCC groups prove the effectiveness of audiovisual materials as a language learning tool, especially if captions are included in the viewing, seeing that students are exposed to contextualised input that offers written support. Hence, considering the benefits for the acquisition of L2 vocabulary, apart from other areas of language, teachers should encourage students to make the most of such entertaining resources both inside and outside the classroom.

7 REFERENCES

Abrahams, Lee B. (2007). Second-Language Reading Comprehension and Vocabulary Learning with Multimedia. *Hispania, 90(1), 98-108.*

Abrams, Zsuzsanna Ittzes. (2014). Using Film to Provide a Context for Teaching L2 Pragmatics. *System, 46, 55-64.*

Barón, Júlia, & Celaya, María Luz. (2022). "May I Do Something For You?": The Effects of Audio-visual Material (Captioned and Non-Captioned) on EFL Pragmatic Learning. *Language Teaching Research, 26(2), 238-255.*

Bisson, Marie-Josée, Van Heuven, Walter J. B., Conklin, Kathy, & Tunney, Richard J. (2014). Processing of Native and Foreign Language Subtitles in Films: An Eye Tracking Study. *Applied Psycholinguistics, 35(2), 399-418.* New York, USA: Cambridge University Press.

Boers, Frank. (2022). Glossing and Vocabulary Learning. *Language Teaching, 55(1), 1-23.* United Kingdom: Cambridge University Press.

Bruti, Silvia. (2016). Audiovisual Texts and Subtitling in the Teaching of Pragmatics. *Trans-kom, 9, 186-207.*

Chang, Anna Ching-Shyang, & Read, John. (2006). The Effects of Listening Support on the Listening Performance of EFL Learners. *TESOL Quarterly, 40(2), 375-397.*

Cheng, Ying-Hsueh, & Good, Robert L. (2009). L1 Glosses: Effects on EFL Learners' Reading Comprehension and Vocabulary Retention. *Reading in a Foreign Language, 21(2), 119-142.*

Choi, Sungmook. (2018). Processing and Learning of Enhanced English Collocations: An Eye Movement Study. *Language Teaching Research, 21(3), 403-426.*

Cintrón-Valentín, Myrna C., & García-Amaya, Lorenzo. (2021). Investigating Textual Enhancement and Captions in L2 Grammar and Vocabulary. *Studies in Second Language Acquisition, 43(5), 1068-1093.*

Collins, Laura, & Muñoz Lahoz, Carme. (2020). *The Foreign Language Classroom: Current Perspectives and Future Considerations.*

Elgort, Irina, Brysbaert, Marc, Stevens, Michaël, & Van Assche, Eva. (2018). Contextual Word Learning During Reading in a Second Language: An Eye-Movement Study. *Studies in Second Language Acquisition, 40(2), 341-366.* New York, USA: Cambridge University Press.

Frumuselu, Anca Daniela, De Maeyer, Sven, Donche, Vincent, & Gutiérrez Colon Plana, María del Mar. (2015). Television Series Inside the EFL Classroom: Bridging the Gap Between Teaching and Learning Informal Language Through Subtitles. *Linguistics and Education,*

32, 107-117. Tarragona, Spain.

García, Boni. (2017). Bilingual Subtitles for Second-Language Acquisition and Application to Engineering Education as Learning Pills. *Computer Applications in Engineering Education, 25(3), 468-479.*

Gass, Susan M. (1988). Integrating Research Areas: A Framework for Second Language Studies. *Applied Linguistics, 9, 198-217.*

Gass, Susan M. (1997). *Input, Interaction, and the Second Language Learner.* Lawrence Erlbaum Associates.

Guillory, Helen G. (2013). The Effects of Keyword Captions to Authentic French Video on Learner Comprehension. *CALICO Journal, 15(1-3), 89-108.*

Hsu, Ching-Kun, Hwang, Gwo-Jen, Chang, Yu-Tzu, & Chang, Chih-Kai. (2013). Effects of Video Caption Modes on English Listening Comprehension and Vocabulary Acquisition Using Handheld Devices. *Educational Technology & Society, 16(1), 403-414.*

Jelani, Nurul Aini Mohd, & Boers, Frank. (2018). Examining Incidental Vocabulary Acquisition from Captioned Video: Does Test Modality Matter? *ITL. Institut Voor Toegepaste Linguistik, 169(1), 169-190.* Canada: University of Western Ontario.

Jewitt, Carey. (2013). Multimodal Methods for Researching Digital Technologies. In Sara Price, Carey Jewitt, & Barry Brown (Eds.), *The SAGE Handbook of Digital Technology Research, 250-265.* California, USA: Sage Publications Limited.

Kaiser, Mark. (2011). New Approaches to Exploiting Film in the Foreign Language Classroom. *L2 Journal 3(2), 232-249.*

Kim, Na Young. (2020). The Effects of the Use of Captions on Low- and High-Level EFL Learners' Speaking Performance. *Linguistic Research, 37, 135-161.*

Ko, Myong Hee. (2012). Glossing and Second Language Vocabulary Learning. *TESOL Quarterly, 46(1), 56-79.*

Krashen, Stephen D. (1985). *The Input Hypothesis: Issues and Implications. 1-32.* London, United Kingdom: Longman Group.

Lee, Minjin, & Révész, Andrea. (2020). Promoting Grammatical Development Through Captions and Textual Enhancement in Multimodal Input-Based Tasks. *Studies in Second Language Acquisition, 42(3), 625-651.* New York, USA: Cambridge University Press.

Li, Mingyue. (2016). *Investigation into the Differential Effects of Subtitles (First Language, Second Language, and Bilingual) on Second Language Vocabulary Acquisition* (Unpublished Doctoral Dissertation). Edinburgh, UK: The University of Edinburgh.

Lin, Phoebe M. S., & Siyanova-Chanturia, Anna. (2015). Internet Television for L2 Vocabulary

Learning. In D. Nunan & J. C. Richards (Eds), *Language Learning Beyond the Classroom, 149-158.* London: Routledge.

Majuddin, Elvenna, Siyanova-Chanturia, Anna, & Boers, Frank. (2021). Incidental Acquisition of Multiword Expressions Through Audiovisual Materials. *Studies in Second Language Acquisition, 43(5), 985-1008.*

Mayer, Richard E. (2009). Multimedia Learning. Cambridge: Cambridge University Press.

Mayer, Richard E., Lee, Hyunjeong, & Peebles, Alanna. (2014). Multimedia Learning in a Second Language: A Cognitive Load Perspective. *Applied Cognitive Psychology, 28(5), 653-660.*

Montero Perez, Maribel. (2022). Second or Foreign Language Learning Through Watching Audiovisual Input and the Role of On-screen Text. *Language Teaching, 55, 163-192.* United Kingdom: Cambridge University Press.

Montero Perez, Maribel, Van Den Noortgate, Wim, & Desmet, Piet Elsevier. (2013). Captioned Video for L2 Listening and Vocabulary Learning: A Meta-Analysis. *System, 41(3), 720-739.*

Muñoz, Carmen. (2022). Audiovisual Input in L2 Learning. *LIA: Language, Interaction and Acquisition, 13(1), 125-143.*

Paivio, Allan. (1986). *Mental Representation: A Dual Coding Approach.* Oxford, UK: Oxford University Press.

Pattemore, Anastasia, & Muñoz, Carmen. (2020). Learning L2 Constructions from Captioned Audio-Visual Exposure: The Effect of Learner-Related Factors. *System (Linköping), 93, 1-13.*

Pemberton, Lyn, Fallahkhair, Sanaz, & Masthoff, Judith. (2004). Towards a Theoretical Framework for Informal Language Learning Via Interactive Television. *In IADIS International Conference Cognition and Exploratory Learning in Digital Age, CELDA, 27-34.*

Peters, Elke. (2019). The Effect of Imagery and On-Screen Text on Foreign Language Vocabulary Learning from Audiovisual Input. *TESOL Quarterly, 53, 1008-1032.*

Pujadas, Geòrgia, & Muñoz, Carmen. (2019). Extensive Viewing of Captioned and Subtitled TV Series: A Study of L2 Vocabulary Learning by Adolescents. *The Language Learning Journal, 47(4), 479-496.*

Rodgers, Michael P. H., & Webb, Stuart. (2017). The Effects of Captions on EFL Learners' Comprehension of English-Language Television Programs. *CALICO Journal, 34(1), 20-38.*

Schlesinger, Izchak M. (1968). *The Syntactic Constituent as a Unit of Decoding: Sentence Structure and the Reading Process.* The Hague: Mouton.

Schmidt, Richard, & Frota, Sylvia Nagem. (1986). Developing Basic Conversational Ability in a Second Language: A Case Study of an Adult Learner of Portuguese. In R. R. Day (Ed.), *Talking to Learn: Conversation in Second Language Acquisition, 237-326.* Rowley, MA: Newbury House.

Sockett, Geoffrey. (2012). From the Cultural Hegemony of English to Online Informal Learning: Cluster Frequency as an Indicator of Relevance in Authentic Documents. *ASp, la revue du GERAS, 60, 5-20.*

Suárez, Maria del Mar. (2018). *Captions and Subtitles. What Would We Do Without Them?*

Sweller, John. (2005). Implications of Cognitive Load Theory for Multimedia Learning. In Richard E. Mayer (Ed.), *The Cambridge Handbook of Multimedia Learning, 19-30.* Cambridge: Cambridge University Press.

Vanderplank, Robert. (1988). The Value of Teletext Sub-titles in Language Learning. *English Language Teaching Journal, 42(4), 272-281.*

Van Lommel, Sven, Laenen, Annouschka, & d'Ydewalle, Géry. (2006). Foreign Grammar Acquisition While Watching Subtitled Television Programs. *British Journal of Educational Psychology, 76, 243-258.*

Wang, Andi, & Pellicer-Sánchez, Ana. (2022). Incidental Vocabulary Learning from Bilingual Subtitled Viewing: An Eye-Tracking Study. *Language Learning, 72(3), 765-805.*

Winke, Paula, Gass, Susan, & Sydorenko, Tetyana. (2010). The Effects of Captioning Videos Used for Foreign Language Listening Activities. *Language Learning & Technology, 14(1), 65-86.*

8 APPENDICES

8.1 Appendix 1: Questionnaire

Hello! I am Maria, a 4th year English Studies undergraduate, and I am conducting a study on the role of audiovisual materials in the acquisition of English as a foreign language as part of my end-of-degree project. Previous to the study, I would appreciate it if you could answer this questionnaire to gather information on your sociolinguistic background; the information will remain anonymous. Thank you for your collaboration.

*Note: The asterisk (*) marks required fields; that is, the questions marked are mandatory.

➢ Please, provide a valid email address.*

➢ Please, write your name and age.*

➢ What is your level of studies?*

- University Postgraduate (Masters or above)

- University Undergraduate (Degree)

- Baccalaureate

- Secondary Education

➢ What is/are your mother tongue(s)?*

➢ What other languages do you speak?*

➢ Do you have any language certificate?*

➢ How many hours per week are you exposed to English?*

➢ Have you ever lived abroad?*

- Yes

- No

➢ If you have, where and for how long?

➢ Have you ever been to an English-speaking country?*

- Yes

- No

➢ If you have, what was the purpose of your visit?

- Leisure

- Study

- Learn English

- Visit friends/family

- Other

- ➤ What streaming platforms do you have?*
  - ▪ Netflix
  - ▪ Prime video
  - ▪ HBO Max
  - ▪ Disney+
  - ▪ Hulu
  - ▪ Other

- ➤ How often do you watch films/series?*
  - ▪ Every day
  - ▪ 4-6 days/week
  - ▪ 1-3 days/week
  - ▪ Once every few weeks
  - ▪ Never

- ➤ Do you watch films/series in their original language?*
  - ▪ Yes
  - ▪ No

- ➤ If you do, what is the original language?
  - ▪ English
  - ▪ Spanish
  - ▪ Catalan
  - ▪ Other

- ➤ If you don't, what is the language of translation?
  - ▪ Spanish
  - ▪ Catalan
  - ▪ Other

- ➤ If you watch films/series in their original language, do you use subtitles?
  - ▪ Yes
  - ▪ No

- ➤ If you do, what is the language of subtitles when you watch English films/series?
  - ▪ English (original language)
  - ▪ Spanish

- Catalan

➢ Do you think the use of subtitles aids in English learning?*

- Yes

- No

- Other

➢ What English variety do you rather watch films/series in?

- British English

- American English

- I have no preference

- Other

➢ What series have you watched from the list below?*

- *Euphoria*

- *The Crown*

- *You*

- *The Bold Type*

- *Red Rose*

- *Manifest*

➢ What films have you watched from the list below?*

- *Joker*

- *Call Me By Your Name*

- *A Star Is Born*

- *Parasite*

- *Knives Out*

- *Oceans' 8*

8.2 Appendix 2: Consent form

Jo, _____, dono el meu consentiment i accepto col·laborar a l'estudi de Maria Port pel seu TFG del grau Estudis Anglesos de la Universitat de Barcelona. Se m'ha informat i entenc que totes les dades són anònimes i que aquest estudi no afectarà la meva nota de classe.

Firmat,   En data __/__/2023

Good evening! As you know, I am conducting a study on the role of audiovisual materials in the acquisition of English as a foreign language as part of my end-of-degree project. You will have to answer 15 questions on English vocabulary; please, if you do not know the answer, choose/write "I don't know", but do not guess the answers. Please, do not look up the meaning of the expressions you do not know. Thank you for your collaboration.

*Note: The asterisk (*) marks required fields; that is, the questions marked are mandatory.

- Give a synonym for *I'm knackered*.*
- What is the meaning of *kick off*?*
  - Leave a place
  - Steal something and run away
  - Start a fight
  - I don't know
- What is the meaning of *tight* in *Don't be tight*?*
  - Strict
  - Mean
  - Cheap
  - I don't know
- Explain in English the meaning of *Don't pull on that thread*.*
- What does *Take the Mickey out of someone* mean?*
  - Make fun of someone
  - Make someone boring
  - Take someone's alcohol
  - I don't know
- What does *hiya* mean?*
- What is the meaning of *Do one*?*
  - Make out
  - Go away
  - Beat up
  - I don't know
- Give a synonym for *tenner*.*
- What does *summat* mean in *I'll lend you summat*?*

- Select the meaning of *proper* in *You're a proper git.**

  - Genuine

  - Absolute

  - Relevant

  - I don't know

- Explain in English the meaning of *Kicking someone's head in.**

- What is the translation of *youse* in Spanish/Catalan?*

- What is the meaning of *scrote* in *Rather stand with a scrote than a slaggy scrubber*?*

  - Worthless person

  - Arrogant person

  - Cruel person

  - I don't know

- What is a *slaggy scrubber*?*

  - Aggressive woman

  - Jealous woman

  - Promiscuous woman

  - I don't know

- What does *fluff* mean in *Did you just fluff?*?*

  - Chuckle

  - Fart

  - Hesitate

  - I don't know

8.4 Appendix 4: Pre-test & Post-test key

Hello! As you know, I am conducting a study on the role of audiovisual materials in the acquisition of English as a foreign language as part of my end-of-degree project. You will have to answer 15 questions on English vocabulary; please, if you do not know the answer, choose/write "I don't know", but do not guess the answers. Thank you for your collaboration.

*Note: The asterisk (*) marks required fields; that is, the questions marked are mandatory.

- Give a synonym for *I'm knackered*.* (DISTRACTOR)

  <u>I'm exhausted</u>

- What is the meaning of *kick off*?*

  - Leave a place

  - Steal something and run away

  - <u>Start a fight</u>

  - I don't know

- What is the meaning of *tight* in *Don't be tight*?*

  - Strict

  - <u>Mean</u>

  - Cheap

  - I don't know

- Explain in English the meaning of *Don't pull on that thread*.*

  <u>Leave an issue alone to avoid making matters worse</u>

- What does *Take the Mickey out of someone* mean?* (DISTRACTOR)

  - <u>Make fun of someone</u>

  - Make someone boring

  - Take someone's alcohol

  - I don't know

- What does *hiya* mean?*

  <u>Hey, how are you?</u>

- What is the meaning of *Do one*?*

  - Make out

  - <u>Go away</u>

  - Beat up

  - I don't know

- Give a synonym for *tenner*.*

  <u>Ten-pound note</u>

- What does *summat* mean in *I'll lend you summat*?*

  <u>Something</u>

- Select the meaning of *proper* in *You're a proper git*.* (DISTRACTOR)

  - Genuine
  - <u>Absolute</u>
  - Relevant
  - I don't know

- Explain the meaning of *Kicking someone's head in*.*

  <u>To hurt or injure somebody by punching them</u>

- What is the translation of *youse* in Spanish/Catalan?*

  <u>Vosotros (Spanish) or vosaltres (Catalan)</u>

- What is the meaning of *scrote* in *Rather stand with a scrote than a slaggy scrubber*?*

  - <u>Worthless person</u>
  - Arrogant person
  - Cruel person
  - I don't know

- What is a *slaggy scrubber*?*

  - Aggressive woman
  - Jealous woman
  - <u>Promiscuous woman</u>
  - I don't know

- What does *fluff* mean in *Did you just fluff?*?* (DISTRACTOR)

  - Chuckle
  - <u>Fart</u>
  - Hesitate
  - I don't know