

# *When to switch captions off? Exploring the effects of L2 proficiency and vocabulary knowledge on comprehension of captioned and uncaptioned TV*

Geòrgia Pujadas ✉

University of Barcelona, Spain

<https://orcid.org/0000-0002-0290-1158>

[georgia.pujadas@ub.edu](mailto:georgia.pujadas@ub.edu)

Carmen Muñoz

University of Barcelona, Spain

<https://orcid.org/0000-0002-7001-4155>

[munoz@ub.edu](mailto:munoz@ub.edu)

### Abstract

The extent to which L2 television is viewed by foreign language learners will depend on the degree to which it is understood. The addition of captions has been shown to support comprehension (e.g., Birulés-Muntané & Soto-Faraco, 2016; Montero-Perez, Peters, & Desmet, 2014), especially when proficiency is low (e.g., Lavour & Bairstow, 2011). Yet, little is known about the extent to which captions benefit comprehension as L2 proficiency increases. This study seeks to investigate the effect of captions at different proficiency levels, and to identify the level at which captions cease to enhance comprehension. A total of 250 Catalan/Spanish university students, who had L2 English proficiency ranging from A1 to C2, viewed nine episodes of an English TV series with and without captions. Results showed that captioned viewing had a significant advantage over uncaptioned viewing in comprehension tests with multiple-choice and true-false items, and that learners with higher L2 proficiency and larger vocabulary performed better. While having access to captions increased the odds of a correct

response independently of learners' L2 proficiency and vocabulary knowledge, the additive benefits of captions were no longer significant at the C2 level, suggesting a threshold beyond which uncaptioned viewing does not negatively impact comprehension. Pedagogical implications are discussed.

*Keywords:* audio-visual input; captions; comprehension; L2 proficiency; vocabulary knowledge

## 1. Introduction

Increasing the amount of exposure to authentic, contextualized input is a key factor in language learning, all the more so in the development of listening skills (Vandergrift, 2007), especially in contexts where foreign language (FL) learners receive little exposure to the FL in everyday life. Streaming platforms such as Netflix, which have increased their number of subscribers worldwide by a factor of almost seven in a decade (Statista, 2024), now provide quick access to a wide range of TV programs in original version and make it easier for people to watch full seasons of their favorite TV series, either through traditional TV sets, laptops, or mobile devices. Since TV series are not originally intended for language learning, however, they may be too difficult to comprehend for learners with less advanced second language (L2) knowledge. In this context, information processing can be aided by switching on captions,<sup>1</sup> which are nowadays widely available in streaming services.

Two findings are reported by most studies on audio-visual comprehension: First, having access to on-screen text (in the first language [L1] or the second language [L2]) has a significant positive effect on comprehension rates (e.g., Montero-Perez, Peters, & Desmet, 2014; Rodgers & Webb, 2017; Winke et al., 2010); and second, L2 knowledge plays a key role in understanding this type of input, with higher proficiency and larger vocabulary knowledge leading to higher comprehension scores (e.g., Montero-Perez, Peters, et al., 2014; Pujadas & Muñoz, 2020). If the desirable outcome for FL teachers and learners is that learners are able to transfer their listening skills to an L2 environment, it seems logical to assume that the goal would be to switch captions off when L2 skills are advanced enough to do so without negatively impacting comprehension. This poses an interesting question for TV viewing in and outside the classroom and addresses a concern among teachers and learners as to the proficiency level that is most appropriate for uncaptioned viewing.

---

<sup>1</sup> In the present paper, the term *captions* (or *L2 captions*) is used to refer to on-screen text in the L2, while the term *subtitles* (or *L1 subtitles*) is used for on-screen text in the learners' L1.

The present study investigates the comprehension of captioned and uncaptioned TV episodes by English as a foreign language (EFL) learners in Spain whose L2 proficiency ranges from beginner to advanced in order to determine the proficiency level at which captions cease to significantly support their comprehension of audio-visual content.

## 2. Literature review

The following sections focus on studies that have addressed the benefits of watching L2 television with or without captions, and on relevant learner- and instruction-related factors.

### 2.1. Switching captions on

The extent to which FL learners will autonomously watch L2 television, which is the end goal of introducing extensive viewing in the classroom (Webb, 2015), depends on the degree to which it is understood. TV programs, which present non-adapted, natural samples of speech that resemble real life, may pose a challenge for FL learners, whose input processing is not as efficient as that of first language listeners (Vandergrift, 2007). Thus, they cannot automatically decode the fast-paced aural stream in order to achieve satisfactory comprehension (e.g., Webb & Rodgers, 2009), and “a large proportion of what they hear may be lost, given the speed of speech and the inability of working memory to process all the information within the time limitations” (Vandergrift, 2007, p. 193). A simple way to help learners decode the message is by switching on captions, that is, subtitles in the language of the audio.

Captions can bridge the gap between learners’ proficiency level and content that lies beyond their L2 skills (Guillory, 1998). Captions encourage the development of segmentation skills (bottom-up processing) and allow students to “break down the input stream and parse it into meaningful constituent structures (Ellis, 2005; Harrington, 2001)” (Winke et al., 2013, p. 255). The opportunity to visualize the speech stream and identify word boundaries aids in the process of word decoding and recognition of their written forms (e.g., Bird & Williams, 2002; Sydorenko, 2010). Captions can also alleviate stream chunking difficulties (Graham, 2006), while reducing learners’ decoding efforts (Buck, 2001).

Research on comprehension of audio-visual input including a captioned versus uncaptioned condition has shown an overall positive effect of captions, with captioned viewing consistently outperforming uncaptioned viewing. Most

studies have found that the advantage is significant (Birulés-Muntaner & Soto-Faraco, 2016; Chung, 1999; Guillory, 1998; Hayati & Mohmedi, 2011; Huang & Eskey, 1999; Latifi et al., 2011; Li, 2014; Markham et al., 2001; Montero-Perez et al., 2014), although others report that the difference does not reach statistical significance (Hsieh, 2020; Lavaur & Bairstow, 2011; Matiolo et al., 2017; Montero-Perez, Peters, et al., 2014; Rodgers & Webb, 2017; Wang & Pellicer-Sánchez, 2022), and one study reports a non-significant advantage when viewing uncaptioned videos (Lee et al., 2021).

However, the benefits from the addition of captions vary widely depending on the study, with differences between captioned and uncaptioned viewing ranging from as little as 2.6% (Rodgers & Webb, 2017) to as much as 46.8% (Li, 2014). This variation is likely due to differences in the design of the studies: The length of videos ranges from 2-minute fragments (e.g., Markham et al., 2001) to full-length 60-minute episodes (Birulés-Muntané & Soto-Faraco, 2016), and the number of items ranges from 8 (Matiolo et al., 2017) to 742 (Rodgers & Webb, 2017). The ratio of the number of items to the length of video might also explain why differences between viewing conditions are (non-)significant. Studies with a higher number of items per minute tend to find non-significant differences between groups (e.g., Montero-Perez et al., 2014: 41 items \* 9 minutes = 4.5 items/minute), while studies with fewer items tend to find larger (significant) differences (e.g., Birulés-Muntané & Soto-Faraco, 2016: 8 items \* 60 minutes = 0.1 items/minute). Despite the overall positive effect of captions on comprehension, it is not yet clear why some learners benefit from captions more than others, but the research points to a key factor: L2 knowledge.

## 2.2. L2 proficiency and vocabulary knowledge

Research has shown that proficiency and vocabulary knowledge play a key role in L2 learning through audio-visual input (e.g., Montero-Perez et al., 2014; Peters, 2019). This is not unexpected, as the accumulation of lexical knowledge facilitates learners' understanding of the input, allows them to focus their attention on novel words, and reduces the learning burden of those unknown words (Webb & Nation, 2017).

In spite of the prominence of the factor, however, only six studies have incorporated learners' vocabulary knowledge (or proficiency level) as an explanatory factor for the comprehension of audio-visual input, with all reporting that L2 knowledge is a strong predictor of comprehension, whether it was a measure of receptive vocabulary knowledge (Montero-Perez, Peters, et al., 2014; Montero-Perez et al., 2014; Pujadas & Muñoz, 2020; Rodgers, 2013; Wang & Pellicer-Sánchez,

2022), a general measure of proficiency (Pujadas & Muñoz, 2020), or the grade from the language course that participants completed (Markham & Peter, 2003). Two studies including advanced EFL learners have also indicated that the benefits from captions may be dependent on the learners' proficiency. Bianchi and Ciabattini (2008) assessed the comprehension of two movie fragments by beginner, intermediate and advanced EFL students. While no final conclusions can be drawn because of the lack of inferential statistics, Bianchi and Ciabattini (2008) reported that there were no differences between captioned and uncaptioned conditions at intermediate level, whereas captions better aided comprehension for beginner and advanced students. Lavour and Bairstow (2011) assessed comprehension with L1 subtitles, L2 captions and no text among beginner, intermediate and advanced EFL learners. Results showed that beginner and intermediate learners performed better with captions (32.8% vs. 22.8% and 55.7% vs. 54.8%, respectively), but learners in the advanced group performed best without on-screen text (74.8% vs. 85.7%).

### 2.3. Switching captions off

Considering the benefits for L2 learning, the question that arises is why FL teachers and learners would choose to switch captions off. First, while captions can reconcile a disparity between the learners' L2 knowledge and the level of difficulty of the audio-visual input, it seems logical to assume that as proficiency increases and the gap narrows, captions could be switched off without a negative impact on comprehension. Second, if the goal is to be able to understand spoken discourse without aid, uncaptioned viewing could be seen as the next desirable step. TV series, which feature lifelike dialogues and everyday situations, can provide a sort of preparatory practice prior to a real communicative situation (such as in study abroad; see Muñoz et al., 2023).

Because studies on comprehension have focused primarily on participants with an intermediate L2 proficiency level – which is a very vague term that refers loosely to a wide range of proficiency – we can only hypothesize what may happen at more advanced levels. While lower-level learners deem captions essential for comprehension, more advanced learners tend to see them only “as a backup to their listening activity” (Pujolà, 2002, p. 254). The same tendency was reported by Chung (1999) in a study combining captions and advanced organizers, where it was found that more advanced students might not need both forms of support. Lavour and Bairstow (2011) also reported that advanced learners found on-screen text to be distracting and they argued that on-screen text, when unnecessary, could disrupt information processing.

Familiarity with the use of captions (and L2 television) may also affect the extent to which learners feel that they need on-screen text. Vanderplank (2019) investigated whether motivation, viewing strategies and attitudes changed after viewing captioned and uncaptioned videos for 6 and 12 weeks. Participants had control over the viewing conditions, which included the option to switch captions on (and off). Data revealed that participants paid less attention to captions as they kept on watching movies, and they grew more confident watching uncaptioned material. This would suggest that students could need captions less over time, and that they may be get a sense of achievement from uncaptioned viewing. Pujadas (2019) investigated changes in beginner, adolescent EFL learners' viewing habits at home after an eight-month classroom intervention involving TV watching and found that students – particularly the ones with higher L2 proficiency – were shifting from L1 subtitles to L2 captions, and then to no text. Results suggested that, with increased familiarity with viewing the original version, the need for on-screen text could dwindle.

#### 2.4. Other factors affecting comprehension in an instructed setting

While the ultimate objective of using audio-visual input in an instructed setting is to promote autonomous extensive viewing at home, in the classroom context, EFL teachers may choose to combine videos with other activities to provide additional semantic support and assist in information processing (Winke et al., 2013). One way of doing so is to present potentially novel words before watching any episodes, using glossaries, or engaging in short language-focused activities, which have been shown to be beneficial for vocabulary learning (e.g., Pujadas & Muñoz, 2019) – ultimately a key goal in the EFL classroom and an efficient way to support comprehension (Webb, 2010). At the same time, however, pre-directing learners' attention to words might deplete their limited attentional resources, and negatively affect content comprehension (e.g., Chang & Read, 2006; Pujadas & Muñoz, 2020).

Studies including vocabulary-related activities prior to viewing have yielded inconsistent results regarding their effect on comprehension. For example, in a study on video viewing comparing three types of advanced organizers (i.e., main characters, vocabulary, and main characters + vocabulary) with 160 low-intermediate learners, Chung and Huang (1998) found that pre-teaching vocabulary was the most efficient strategy to aid comprehension. On the other hand, in a study investigating different forms of assisting in listening comprehension with 160 low-intermediate university-level learners, Chang and Read (2006) found that studying topic-related vocabulary was the least effective type of support compared to previewing the test questions, repeating the input, or providing information on the topic. Pujadas and Muñoz (2020) found in a study with 106

adolescent learners from the A1 to B2 level that participants who were explicitly taught target items tended to have lower comprehension, even though they were the ones who obtained higher vocabulary gains. Pujadas and Muñoz suggest that learners – at that particular age and proficiency level – may have found it difficult to divide their attentional resources between the two tasks (VanPatten, 2002), but such additional vocabulary-focused activities may not be attention-depleting for older or more advanced learners.

## 2.5. Aim and research questions

The main aim of the present study is to investigate the effect of captions on viewing comprehension at different proficiency levels and to potentially identify the proficiency level at which the additive effects of captions are no longer significant. Though the logical prediction is that captions will be less necessary as learners become more advanced, the precise proficiency threshold for satisfactory uncaptioned viewing is still unknown. The analyses include two learner-related variables (proficiency and vocabulary knowledge), as well as learning context and test-related variables (i.e., pre-teaching of vocabulary, item format and type of information). The research questions (RQs) are formulated as follows:

- RQ1: To what extent do captions aid the comprehension of TV series?
- RQ2: To what extent is the potential effect of captions affected by learners' L2 proficiency and vocabulary knowledge?
- RQ3: At what proficiency level do captions cease to significantly support comprehension of audio-visual content?

## 3. Methods

### 3.1. Participants

The initial sample of participants were 352 second-year university students from three different BA programs who were enrolled in a compulsory English course. Second-year classes were targeted because they typically include a wider range of L2 proficiency levels, compared to more advanced courses. Data were collected from the entire population available (Brysbart & Stevens, 2018). From the initial sample, only those who had 85% attendance or more and had signed the consent form were included in the final sample, leaving a total of 250 participants (180 females, 70 males). Four intact classes were assigned to the captioned group (CG,

$N = 159$ ) and four to the uncaptioned group (UG,  $N = 91$ ). Two classes in each group were also pre-taught vocabulary items appearing in the episodes.

Participants were Catalan-Spanish balanced bilinguals, with a mean proficiency level of B2/C1 in English, ranging from A1 (beginner level) to C2 (proficient level). Participants had a mean vocabulary knowledge of around 5000 words – ranging from 1900 to 8300 words – as measured by the X\_Lex and Y\_Lex tests (Meara & Milton, 2003; Meara & Miralpeix, 2006).

### 3.2. Audio-visual materials

Nine consecutive episodes from the 5th season of the series *I Love Lucy* (Oppenheimer & Arnaz, 1951) were selected for the study. The series was chosen because it had neither been recently nor currently broadcast, and its original release date made it fairly unknown to participants. Secondly, while the series had a season-long story arc, each episode contained a full story arc itself, which allowed viewers to gather background information while being able to follow the episodes even if they skipped one throughout the term.

The episodes had a mean running time of about 30 minutes – including opening credits – which added up to a total viewing time of 270 minutes and a total of 32,374 tokens by the end of the study. The scripts were analyzed through the RANGE software (Nation & Heatley, 2002), which showed that – on average – the episodes reached 92.16% coverage at the 1,000-word level plus proper nouns, marginal words, and words uttered in Spanish,<sup>2</sup> 95.68% at the 2,000-word level, and 97.39% at the 3,000-word level. As for the individual episodes, seven out of nine reached 95% coverage at the 2,000-word level, and two out of nine reached 98% at the 3,000-word level.

### 3.3. L2 proficiency and vocabulary knowledge tests

Participants' initial proficiency level (henceforth *general proficiency*) was measured through the Oxford Placement Test (OPT). The OPT test was considered appropriate because it includes both a grammar and a listening section, and listening skills are notably relevant in this learning context. The OPT has also been calibrated against the system of levels provided by the *Common European Framework of Reference* (CEFR) and allows to reliably place test-takers into their corresponding CEFR levels.

---

<sup>2</sup> Since one of the main characters was of Cuban origin, there was a small percentage of words (0.86%) in Spanish, which were considered known by participants together with proper nouns and marginal words.



Vocabulary knowledge was measured using the X\_Lex test (Meara & Milton, 2003) and Y\_Lex test (Meara & Miralpeix, 2006), which provide an estimate of learners' receptive vocabulary knowledge of the most frequent 5,000 and 10,000 word families respectively. Each test presents 20 words from each 1,000-word family in a *yes-no* format (alongside non-words to correct for guessing), and the total score is calculated by adding up the words known in each of the 10,000-word lists. The scores on both tests were added to obtain a single measure after consultation with one of the test creators (Miralpeix, personal communication, August 27, 2022).

### 3.4. Comprehension tests

Comprehension was assessed after each of the nine episodes through post-viewing tests. Tests were administered in the learners' L1 (Catalan or Spanish) to avoid errors ascribable to poor comprehension of the items (Vandergrift, 2007). Each test consisted of 10 questions, which included 5 multiple-choice items (MC) – with three possible answers (the correct one and two distractors) – and 5 true-false items (TF). Participants were advised to leave the question blank if they were completely unsure of the answer – to mitigate the effect of guessing – and only guess if they had enough clues to make an informed choice, capturing partial knowledge. The questions also included two types of information retrieval: textually explicit items (TE) (the answer was explicitly stated in the episode) and inferential items (IN) (learners had to put the answer together or deduce it, integrating pieces of information scattered across the episode to figure out the central gist or idea). Items were scored as either correct (= 1) or incorrect (= 0).

This operationalization follows that of Pujadas and Muñoz's (2020) study (see also Alptekin & Erçetin, 2010; Davey & McBride, 1986; Rodgers, 2013). Since prior research has reported that comprehension scores can be affected by the response format (e.g., Cheng, 2004) and the type of information elicited by the questions (e.g., Pujadas & Muñoz, 2020; Rodgers, 2018), both variables were factored in the analysis.

### 3.5. Procedure

The study took place over ten weeks. Proficiency and vocabulary tests were administered by the first author, while the nine 50-minute-long viewing sessions were carried out by the course teachers. The two groups with vocabulary instruction started the session with a short, 5-minute vocabulary pre-task.<sup>3</sup> Participants

---

<sup>3</sup> Vocabulary pre-tasks (or "vocabulary training" tasks) were short activities focused on teaching five target words in each episode. The tasks were presented in English, and included word-searches (2

then viewed the episode (captioned or uncaptioned) and completed the comprehension test. The two groups without vocabulary instruction started the session directly by viewing the episode (captioned or uncaptioned) (see Figure 1).

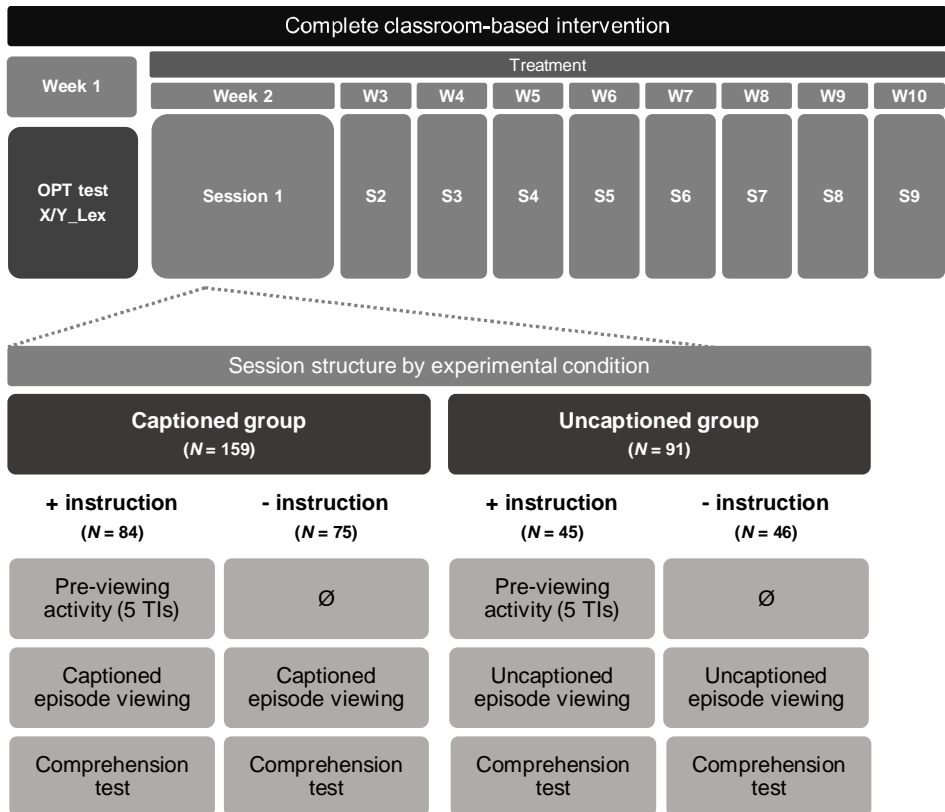


Figure 1 Outline of the classroom-based intervention

### 3.6. Data analysis

For the first research question, the total number of correct responses for participants in the captioned (CG) and uncaptioned group (UG) was added up, and a Welch's ANOVA (a Levene's test showed variances were unequal) was run to compare each viewing condition. The maximum score was 90 (9 tests \* 10 items), but 35 out of the 250 participants missed one or two sessions, and thus the analysis included 22,030 valid cases.

episodes), fill-in-the-blanks (4), matching exercises (2) and crosswords (1). Activities were completed individually, corrected orally by the teacher, and then collected before viewing the episode.

For the second research question, only participants who had completed both the OPT test ( $N = 222$ ) and the vocabulary knowledge test ( $N = 233$ ) were included in the analysis, leaving 210 participants with both tests completed ( $CG = 139$ ;  $UG = 71$ ) and 18,620 valid observations. A generalized linear model (GLM) with repeated measures was calculated with comprehension score (0/1) as the dependent variable, and captions (yes/no), general proficiency (continuous), vocabulary knowledge (continuous), instruction (yes/no), item format (MC/TF) and information type (TE/IN) as fixed factors, including all two-way interactions between the captions condition and the other five variables. The variance inflation factor (VIF) was calculated to check multicollinearity for the two learner-related variables (i.e., general proficiency and vocabulary knowledge). The VIF obtained was 1.68, which indicates an absence of multicollinearity, although a Pearson correlation revealed that there was a significant, moderate correlation between them ( $r = .658$ ,  $p < .001$ ). These two variables were centered (i.e., z-scored) before running the model. All non-significant interactions and main effects ( $p < .10$ ) were removed one by one to arrive at the best fitting model.

For the third research question, proficiency was recoded into CEFR levels according to participants' OPT scores, initially distributing them into 6 proficiency bands, from A1 to C2. Due to the very small sample of participants found in the A1 band, participants in that level were excluded from the following analysis, leaving a total of 212 participants. The A2 and C2 groups were also small compared to the other proficiency bands, but they were kept in the analysis to further explore the effects of captions at the lowest and highest proficiency levels, although the results for these two groups should be interpreted with caution. The C1 group – which represented a large proportion of the sample – was divided into two subgroups of similar size (C1.1 and C1.2) (see Table 5 for the participants' distribution). A second GLM with repeated measures was calculated with the re-categorized proficiency with six levels (A2, B1, B2, C1.1, C1.2, C2). This model was based on 18,770 observations.

## 4. Results

Preliminary descriptive analyses of the participants' proficiency level and vocabulary knowledge are presented first, followed by the inferential statistical analyses conducted to answer the research questions.

### 4.1. Preliminary analyses: Proficiency and vocabulary knowledge

Table 1 shows the descriptive statistics for general proficiency and vocabulary knowledge for each viewing condition. Welch's ANOVAs revealed that there were

no significant differences amongst groups in terms of general proficiency ( $F(3, 218) = 1.500$ ;  $p = .220$ ) nor vocabulary knowledge ( $F(3, 229) = 1.420$ ;  $p = .238$ ).

Table 1 Proficiency and vocabulary knowledge by condition, with standard deviation in brackets

	<i>N</i>	General proficiency (max. 200)	<i>N</i>	Vocabulary knowledge (max. 10000)
Captioned Group (CG)	147	143.99 (17.76)	149	5119.46 (1220.02)
Uncaptioned Group (UG)	75	142.28 (21.71)	84	4854.17 (1369.77)
All	222	143.41 (19.16)	233	5023.82 (1279.48)

#### 4.2. Effect of captions on viewing comprehension

Table 2 shows the total number of correct and incorrect responses per each viewing condition (percentage in brackets) across the nine episodes. Participants in the CG obtained 87.4% of correct responses, while in the UG the percentage was 73.7%. This represents a difference of 13.7%, with the CG outperforming the UG. Results from a Welch's ANOVA showed that this difference was significant, with a small effect size ( $F(1, 102) = 123.03$ ,  $p < .001$ ;  $r = 1.39$ ).

Table 2 Total number (and percentage) of correct and incorrect responses with and without captions

	Correct responses	Incorrect responses	Total number of responses
Captioned Group (CG)	12304 (87.4%)	1776 (12.6%)	14080
Uncaptioned Group (UG)	5858 (73.7%)	2092 (26.3%)	7950
All	18162 (82.4%)	3868 (17.6%)	20030

#### 4.3. Effect of proficiency and vocabulary knowledge

A GLM revealed that four factors significantly contributed to the model (captions, general proficiency, vocabulary knowledge, information type) and one contributed marginally (item format). Three significant interactions emerged between captions and the variables of general proficiency, item format and item type. Table 3 shows the main fixed factors effects and interactions, and Table 4 illustrates the significant fixed main effects for categorical variables.

Table 3 Results from the GLM: effects of fixed factors on comprehension scores

Terms	Coeff	SD	T	p	Exp coeff	95% CI for Exp coeff	
						Lower	Upper
Intercept	1.591	.0732	21.749	<.001	4.910	4.254	5.667
Captions (yes)	.314	.0948	3.315	<.001	1.369	1.137	1.649
Captions (no)	0 <sup>a</sup>	.	.	.	.	.	.
Proficiency	.245	.0447	5.475	<.001	1.277	1.170	1.395
Vocabulary	.138	.0378	3.638	<.001	1.148	1.066	1.236
Format (MC)	.112	.0592	1.890	.059	1.118	.996	1.256
Format (TF)	0 <sup>a</sup>	.	.	.	.	.	.
Type (TE)	-.745	.0681	-10.951	<.001	.475	.415	.542
Type (IN)	0 <sup>a</sup>	.	.	.	.	.	.
C(yes) * Prof.	-.151	.0584	-2.591	.010	.860	.767	.964
C(yes) * F (MC)	.435	.0815	5.333	<.001	1.545	1.316	1.812
C(yes) * T (TE)	.434	.0907	4.783	<.001	1.543	1.292	1.843

Note. MC = multiple-choice; TF = true-false; TE = textually explicit; IN = inferential; Coeff = coefficient; Exp coeff = exponential coefficient; CI = confidence interval; <sup>a</sup> Redundant coefficient is set to zero

Table 4 Results from GLM: effects of fixed main effects (categorical variables)

Factor	Group	$\bar{X}$ (SE)	df	F	p
Captions	Yes	.883 (.004)	1, 18611	114.693	< .001
	No	.781 (.008)			
Format	MC	.860 (.005)	1, 18611	66.235	< .001
	TF	.815 (.005)			
Type	TE	.800 (.005)	1, 18611	154.077	< .001
	IN	.871 (.005)			

Note.  $\bar{X}$  = mean; SE = Standard Error; MC = multiple-choice; TF = true-false; TE = textually explicit; IN = inferential

Results showed that the presence of captions had a significant positive effect on comprehension scores ( $p < .001$ ), indicating that an average learner's comprehension scores were 10.2% higher when they had access to captions – all other factors held constant – with participants in the CG having overall 88.3% of correct responses in contrast with the 78.1% in the UG. The odds of selecting the correct response in the comprehension tests were about 1.37 times higher for captioned viewing; in other words, the odds of answering correctly were 37% higher in the CG compared to the UG. An interaction emerged between the captions condition and general proficiency, item format and item type, suggesting that the effect of these three parameters needs to be explained in relation to the presence (or absence) of captions.

The model revealed that both L2 knowledge measures were positively related to comprehension, with higher general proficiency and larger vocabulary knowledge leading to higher scores in comprehension, regardless of the caption condition. When increasing general proficiency by one *SD*, the odds of a correct response increased by 27.7%. The interaction between proficiency and caption condition ( $p = .010$ ) indicated that the effect of general proficiency – while significant in both caption

conditions – was stronger when captions were *not* present. This interaction was further explored through a regression analysis, which showed that the score in the general proficiency test accounted for 56.6% of the variance in comprehension in the UG, while it accounted for 10.4% of the variance in the CG (see Figure 2). In the case of vocabulary knowledge, the odds of a correct response increased by 14.8% when vocabulary knowledge increased by one *SD*. The percentage of variance accounted for by this measure was of 20.3% in the UG, and 18.8% in the CG (see Figure 3).

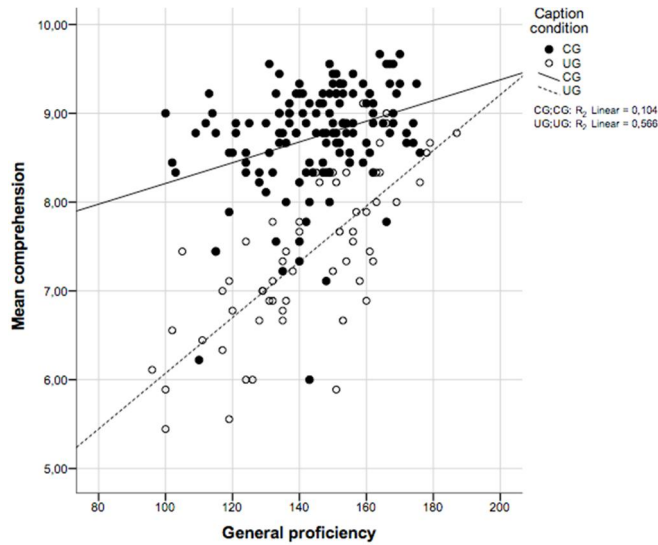


Figure 2 Interaction between captions and general proficiency

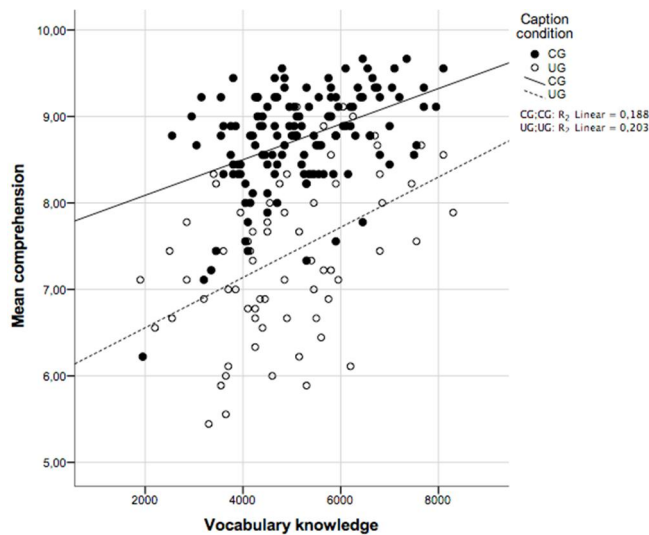


Figure 3 Interaction between captions and vocabulary knowledge

Results also indicated that vocabulary instruction had no effect on comprehension rates and that the groups that had been pre-taught target items prior to viewing had similar comprehension rates than the groups that were not receiving instruction (i.e., 83% vs. 82%, respectively). Regarding the two test-related variables, the model showed that both item format and item type predicted comprehension rates, although both interacted significantly with the caption condition. In the case of item format, a marginally significant main effect was found ( $p = .059$ ), with MC items receiving a higher number of correct responses (85.6%) compared to TF items (81.8%). An interaction revealed that the effect of this variable was stronger in the CG (MC 90.9% vs. TF 85.2%) ( $F(1, 18611) = 98.230, p < .001$ ) and only marginal in the UG (MC 79.1% vs. TF 77.2%) ( $F(1, 18611) = 3.585, p = .058$ ). Type of information was also significantly correlated with comprehension ( $p < .001$ ), independently of the caption condition, with IN items receiving a higher number of correct responses (87.1%) than TE items (80%). An interaction showed, however, that the effect of this variable was stronger in the UG, where the difference was 12.7% between IN (83.8%) and TE items (71.1%), as opposed to the smaller (yet still significant) difference of 3.2% in the CG (IN 89.8% vs. TE 86.6%).

#### 4.4. Proficiency threshold for comprehension without captions

A second GLM showed that the CEFR level was significantly related to comprehension scores ( $F(5, 18758) = 13.194, p < .001$ ), with higher proficiency leading to higher comprehension scores. Pairwise comparisons revealed that differences were not significant amongst the A2, B1 and B2 levels, nor amongst the C1.1, C1.2 and C2 levels, but they were significant when comparing the A2, B1, B2 levels (beginner / intermediate) against the C1.1, C1.2, C2 levels (advanced) (ranging from  $p < .001$  to  $p = .006$ ).

In contrast with the first model, however, no interaction emerged between the captions condition and the CEFR bands ( $F(5, 18758) = 1.156, p = .328$ ), implying that the effect of this variable was independent of the captions condition. Table 5 shows the mean comprehension score by caption condition and proficiency level, as well as the contrast estimate and the significance value of each contrast. Pairwise contrasts confirmed that differences between CEFR bands within each caption condition were significant, although differences in comprehension scores due to proficiency were significantly larger for uncaptioned viewing. Regarding differences within each proficiency level, pairwise contrasts revealed that differences between the CG and UG were significant ( $p < .001$ ) in all levels but one: the C2 level ( $F(1, 18758) = 1.527, p = .217$ ) (see Table 5). In other words,

participants in the CG consistently outperformed participants in the UG up to the C2 level, at which point differences were no longer significant – although the CG still had a higher score than the UG. Figure 4 illustrates this interplay between the caption condition and proficiency level. As it can be observed, participants in the CG had a mean comprehension over 80%,<sup>4</sup> independently of the proficiency level, while this percentage could only be achieved at the C1.2 level and the C2 level in the UG.

Table 5 Mean comprehension (in percentage) for CG and UG per each CEFR level

	<i>N</i>	CG Mean (SE)	<i>N</i>	UG Mean (SE)	Contrast (SE)	<i>F</i>	<i>t</i>	<i>p</i>
A2	11	82.5% (1.9)	7	64.9% (3.0)	15.3% (2.2)	23.792	4.878	<.001
B1	22	86.6% (1.2)	16	69.8% (1.9)	15.3% (2.2)	54.258	7.366	<.001
B2	48	85.5% (0.9)	11	72.5% (2.3)	11.2% (2.3)	28.875	5.374	<.001
C1.1	30	89.7% (0.9)	18	77.2% (1.7)	10.7% (1.8)	42.466	6.517	<.001
C1.2	22	90.4% (1.1)	12	81.4% (1.9)	7.6% (2.0)	17.108	4.136	<.001
C2	8	88.9% (1.8)	6	86.6% (2.4)	2.8% (2.7)	1.527	1.236	.217

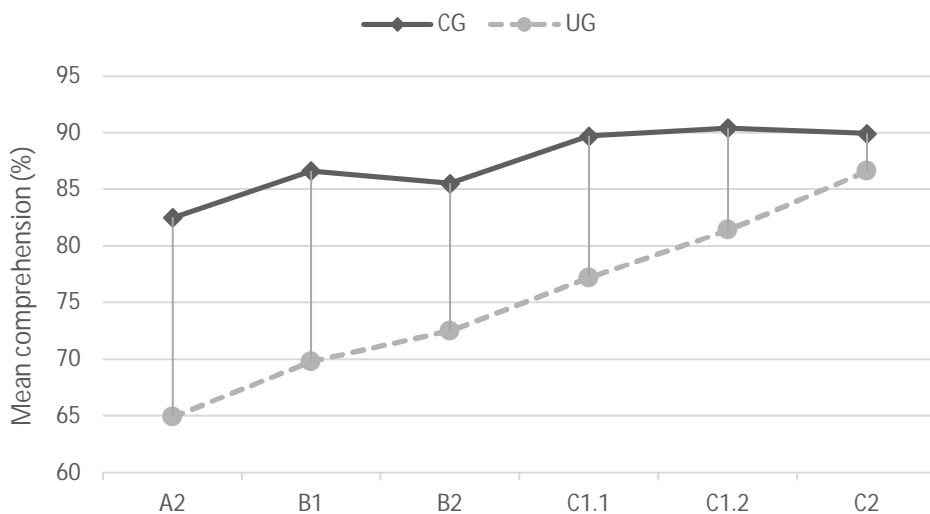


Figure 4 Mean comprehension for CG and UG by CEFR level

<sup>4</sup> Only two studies have suggested an actual definition of “adequate” comprehension. The first one was Laufer (1989), who considered that adequate comprehension was a score of 55%. The other one was Hu and Nation (2000), who suggested a higher number: 85.7% (12 correct items out of 14). This is similar to the percentage of 80% (i.e., 8 out of 10) considered to be the score for satisfactory comprehension in the present study.



## 5. Discussion

This study sought to investigate the effect of captions on viewing comprehension at different proficiency levels. It also sought to explore the extent to which participants' proficiency and vocabulary knowledge might affect comprehension during captioned and uncaptioned viewing, and to identify a tentative proficiency threshold for satisfactory comprehension without captions.

### 5.1. The benefits of captions and their interplay with learners' L2 knowledge

Our first research question looked into the additive effect of captions on comprehension. Overall, the results showed a significant advantage of captioned viewing over uncaptioned viewing for content comprehension, holding all other variables constant. The use of captions increased the odds of a correct answer by 37%, while participants with access to captions scored, on average, 10% higher than those in the uncaptioned group. These results are in line with previous findings from studies comparing captioned and uncaptioned viewing, which have consistently found a significant overall advantage from captions (e.g., Birulés-Muntaner & Soto-Faraco, 2016; Montero-Perez et al., 2014).

The second research question investigated the extent to which L2 proficiency and vocabulary knowledge mediate the potential effect of captions. Results from the GLM showed that both L2 measures predicted comprehension, independently of the caption condition, with higher proficiency and larger vocabulary knowledge leading to higher comprehension rates. The results also concur with prior research in the field, which has found that both variables are good predictors of comprehension (e.g., Markham & Peter, 2003; Montero-Perez et al., 2014; Montero-Perez, Peters, et al., 2014). In the present study, general proficiency emerged as a better predictor than vocabulary knowledge when viewing without captions, as the variable interacted with the caption condition and the effect was considerably stronger when captions were not present (56.6% vs 10.4%). By contrast, vocabulary knowledge accounted for a similar amount of variance in both conditions. This is a noteworthy finding, as it shows that, although correlated, the explanatory power of the two variables may differ. This finding also highlights the importance of including L2 knowledge as a factor in the analysis, rather than merely using it as a descriptive variable.

The presence of vocabulary-focused activities, a relevant aspect within the EFL classroom context, did not significantly affect comprehension rates in either the CG or the UG, with both conditions obtaining almost identical comprehension results. The results contrast with findings obtained by Pujadas and

Muñoz (2020), who found that pre-teaching vocabulary had a small negative impact on comprehension. A plausible explanation might be that participants in the present study were older (i.e., university students) and more proficient, so they were better able to cope with both activities simultaneously.

Regarding item format, it was found that MC items received a slightly higher number of correct responses than TF items, but the difference was only significant when captions were present. This contrasts with findings by Pujadas and Muñoz (2020), who found that TF items had a higher number of correct responses than MC items. It is possible that participants in the current study (who were older and more proficient) could also select the correct option by eliminating the ones that they did not hear (or read in the captions), while younger and less proficient learners may have had difficulties in doing so. The type of information also appeared as a significant predictor of comprehension, with IN items receiving a higher number of correct responses than TE items (especially in the UG), which is partially in line with the results found by Pujadas and Muñoz (2020). TE items required that learners understood details that could go unnoticed owing to the fast-paced speech of the episodes, while for IN items students could put together information from different scenes in the episodes, which might have well compensated for a missed piece of information. This strategy could be followed regardless of whether captions are on or off. For TE items, however, participants with access to captions could *read* the answers (although they were paraphrased), while participants in the UG did not have this option, thus explaining why TE items, in particular, were more challenging in this condition.

## 5.2. Proficiency threshold to switch captions off

Our third research question aimed to explore the proficiency level at which captions could be switched off without negatively impacting comprehension. The results suggested a proficiency threshold at the C2 level, beyond which comprehension scores were not significantly different between the CG and the UG, although the CG still performed better than the UG. This contrasts with findings by Lavaur and Baristow (2011), who reported that differences between caption conditions were significant only in the advanced group and that the uncaptioned group performed significantly better at this proficiency level.

The results of the current study require confirmation with a larger sample but offer a preliminary threshold for satisfactory uncaptioned viewing. For students below the C2 level, uncaptioned viewing had a significant negative effect on comprehension rates. The results also revealed, however, that the gap between the CG and the UG narrowed as proficiency increased, even if the CG

consistently outperformed the UG at all five levels. At the A2 level, participants with access to captions scored 17.6% better than their counterparts, but the difference decreased to 16.8% at the B1, 13% at B2, 12.6% at C1.1 levels, 8.9 % at the C1.2 level, and ultimately to a non-significant 3.7% at the C2 level.

The results also showed that there was a significant difference in comprehension scores between intermediate (i.e., B1, B2) and advanced (i.e., C1, C2) learners, but not within each broad level (i.e., within the B level or within the C level). For uncaptioned viewing, however, a clear tendency could be observed: students at each proficiency level performed better than the level below (see Figure 3), even if not significantly. This suggests that captions could compensate for differences in the level of proficiency (especially if the levels are close) when students view the same audio-visual materials together. It is important to note that these tendencies could not have been observed if participants had been grouped into broader levels (i.e., intermediate, advanced), suggesting the need for a fine-grained description of learners' L2 proficiency when assessing the effects of L2 knowledge.

In the present sample, students who fell in the intermediate category had a mean comprehension score of about 86% with captions and 71% without captions. These percentages were considerably higher than those generally found in other studies with intermediate participants (e.g., 60% and 54% in Montero-Perez, Peters & Desmet, 2014; 66.5% and 64% in Rodgers & Webb, 2017). One reason for the difference might have been the number of items in the comprehension tests (i.e., 10 items for a 20-minute episode), which was smaller than in most studies (e.g., Rodgers & Webb, 2017; Montero-Perez et al., 2014). For example, the study by Matielo et al. (2017), which featured a similar ratio of items per episode length (i.e., 8 items for a 20-minute episode), found similar comprehension percentages to the present study (86.6% vs. 76.6% in specific questions). Although the disparity in study designs, input length and test format make comparisons difficult, the mean comprehension for studies with "intermediate" participants was 70% with captions and 55% without captions, with an average difference of 15% (Birulés-Muntaner & Soto-Faraco, 2016; Chung, 1999; Guillory, 1998; Hayati & Mohmedi, 2011; Hsieh, 2020; Huang & Eskey, 1999; Latifi et al., 2011; Lavour & Bairstow, 2011; Lee et al., 2021; Li, 2014; Markham et al., 2001; Matielo et al., 2017; Montero-Perez et al., 2014; Montero-Perez, Peters, et al., 2014; Rodgers & Webb, 2017). This is similar to the 15% difference in the present study for participants at the B1 and B2 levels. Finally, the data also showed that a satisfactory level of comprehension (e.g., 80%) could generally be achieved independently of proficiency when captions were present and that without caption support this level of comprehension could only be achieved at the most advanced proficiency levels.

This study provides important insights. First, it confirms the advantage of having a higher proficiency level and larger vocabulary knowledge for content comprehension, especially for uncaptioned viewing – the most challenging condition – with more proficient learners obtaining the best results. Second, it provides evidence that, at the highest levels of proficiency, the absence of captions does not significantly impact comprehension, and it therefore tentatively sets a threshold at the proficiency level of C2. FL teachers and learners should be aware, then, that captions are recommended below this level when the objective is content comprehension. Another finding is that the addition of vocabulary activities, which have been shown to be a valuable tool for vocabulary learning using this medium (e.g., Pujadas & Muñoz, 2019), neither hindered nor enhanced comprehension. Finally, the effects of the type of information asked for and its interaction with the presence (or absence) of captions suggest that learners process the two types of information (i.e., textually explicit or inferential) differently. The effect of these variables also highlights the need to take test-related features into account and to include different format types when testing.

## 6. Conclusion

This study confirmed the additive benefits of captions and the advantage of higher proficiency and larger vocabulary knowledge for viewing comprehension (especially for uncaptioned viewing), while it also tentatively sets a threshold at the C2 level for satisfactory comprehension without captions. The present study contributes to the field of second language learning through audio-visual input, and specifically advances our knowledge concerning the effects of captioning on audio-visual comprehension at different proficiency levels. To the best of the authors' knowledge, this is the first study that has sought to set a proficiency threshold for viewing comprehension without captions, laying the groundwork for future research in the area.

The study has several pedagogical implications. First, considering the variety of proficiency levels found in EFL classrooms, captions have been shown to compensate for such differences in L2 skills, allowing learners with a wide range of proficiency levels (from A2 to C2, in the case of the present sample) to follow the content of episodes satisfactorily, with a mean comprehension at above 80% for captioned viewing. Second, the results suggest that more advanced learners can follow that content equally with or without captions. When guiding autonomous viewing at home, EFL instructors could suggest uncaptioned viewing at home when the students' level is advanced enough (i.e., C1.2 or higher). Additionally, the results indicate that, at the university level, audio-visual input can be combined in the classroom context with language-focused activities without compromising comprehension.

The study suffers from a handful of limitations. The C2-level group was considerably smaller than the other groups, so the analyses that included this proficiency group were exploratory and the results should be interpreted with caution. Small sample sizes present higher variability, and replication with a larger advanced-level group would improve the validity of our findings. Our current results, however, provide a first step towards establishing a threshold for satisfactorily viewing TV without captions. Another limitation is that we did not control for random guessing by participants, which might have biased results (especially for TF items); adding an *I don't know* option could help to mitigate the issue (Schmitt, 2010). The results are also contingent on the input (i.e., the TV series) chosen for the study. The variability in language, content and lexical demands across TV series is likely to affect the degree to which they are understood, as research comparing several episodes of a TV series has shown considerable variability in comprehension even within the same TV series (e.g., Rodgers & Webb, 2017). Future research that looks into the interplay between captions and proficiency using series with varying lexical demands could contribute to a better understanding of the variables that mediate viewing comprehension and support EFL teachers and learners in selecting the TV series and viewing conditions that are best suited to learners' proficiency levels.

### Acknowledgments

Grant PID2019-110594GB-I00 from the Spanish Ministry of Science and Innovation. Margarita Salas Grant for the training of young doctors from the Spanish Ministry of Universities.

We would like to thank Dr. Suárez for her help with the data collection, as well as all the students that participated in the study and the professors that conducted the viewing sessions.

## References

- Alptekin, C., & Erçetin, G. (2010). The role of L1 and L2 working memory in literal and inferential comprehension in L2 reading. *Journal of Research in Reading, 33*(2), 206-219. <https://doi.org/10.1111/j.1467-9817.2009.01412.x>
- Bianchi, F., & Ciabattini, T. (2008). Captions and subtitles in EFL learning: An investigative study in a comprehensive computer environment. In Baldry, A., Pavesi, M., & Taylor Torsello, C. (Eds.), *From didactas to ecolingua* (pp. 69-90). Edizioni Università di Trieste.
- Bird, S. A., & Williams, J. N. (2002). The effect of bimodal input on implicit and explicit memory: An investigation into the benefits of within-language subtitling. *Applied Psycholinguistics, 23*(4), 509-533. <https://doi.org/10.1017/S0142716402004022>
- Birulés-Muntané, J., & Soto-Faraco, S. (2016). Watching subtitled films can help learning foreign languages. *PloS One, 11*(6), e0158409. <https://doi.org/10.1371/journal.pone.0158409>
- Brybaert, M., & Stevens, M. (2018). Power analysis and effect size in mixed effects models: A tutorial. *Journal of Cognition, 1*(1), 1-20. <https://doi.org/10.5334/joc.10>
- Buck, G. (2001). *Assessing listening*. Cambridge University Press. <https://doi.org/10.1017/CBO9780511732959>
- Chang, A., & Read, J. (2006). The effects of listening support on the listening performance of EFL learners. *TESOL Quarterly, 40*(2), 375-397. <https://doi.org/10.2307/40264527>
- Cheng, H. (2004). A comparison of multiple-choice and open-ended response formats for the assessment of listening proficiency in English. *Foreign Language Annals, 37*(4), 544-553. <https://doi.org/10.1111/j.1944-9720.2004.tb02421>
- Chung, J. M. (1999). The effects of using video texts supported with advance organizers and captions on Chinese college students' listening comprehension: An empirical study. *Foreign Language Annals, 32*(3), 295-308. <https://doi.org/10.1111/j.1944-9720.1999.tb01342.x>
- Chung, J. M., & Huang, S. C. (1998). The effects of three aural advance organizers for video viewing in a foreign language classroom. *System, 26*(4), 553-565. [https://doi.org/10.1016/S0346-251X\(98\)00037-2](https://doi.org/10.1016/S0346-251X(98)00037-2)
- Davey, B., & McBride, S. (1986). Effects of question-generation training on reading comprehension. *Journal of Educational Psychology, 78*(4), 256-262. <https://doi.org/10.1037/0022-0663.78.4.256>
- Ellis, N. C. (2005). Constructions, chunking, and connectionism: The emergence of second language structure. In C. J. Doughty & M. H. Long (Eds.), *The handbook of second language acquisition* (pp. 63-103). Blackwell.

- Graham, S. (2006). Listening comprehension: The learners' perspective. *System*, 34(2), 165-182. <https://doi.org/10.1016/j.system.2005.11.001>
- Guillory, H. (1998). The effects of keyword captions to authentic French video on learner comprehension. *CALICO Journal*, 15(1-3), 89-108. <https://www.jstor.org/stable/24147690>
- Harrington, M. (2001). Sentence processing. In P. Robinson (Ed.), *Cognition and second language instruction* (pp. 91-124). Cambridge University Press.
- Hayati, A., & Mohmedi, F. (2011). The effect of films with and without subtitles on listening comprehension of EFL learners. *British Journal of Educational Technology*, 42(1), 181-192. <https://doi.org/10.1111/j.1467-8535.2009.01004.x>
- Hsieh, Y. (2020). Effects of video captioning on EFL vocabulary learning and listening comprehension. *Computer Assisted Language Learning*, 33(5-6), 567-589. <https://doi.org/10.1080/09588221.2019.1577898>
- Hu, M., & Nation, P. (2000). Unknown vocabulary density and reading comprehension. *Reading in a Foreign Language*, 13, 403-430.
- Huang, H., & Eskey, D. (1999). The effects of closed-captioned television on the listening comprehension of intermediate English as a second language (ESL) students. *Journal of Educational Technology Systems*, 28(1), 75-96. <https://doi.org/10.2190/RG06-LYWB-216Y-R27>
- Latifi, M., Mobalegh, A., & Mohammadi, E. (2011). Movie subtitles and the improvement of listening comprehension ability: Does it help? *The Journal of Language Learning and Teaching*, 1(2), 18-29.
- Laufer, B. (1989) What percentage of text-lexis is essential for comprehension? In C. Lauren & M. Nordman (Eds.), *Special language: From human thinking to thinking machines* (pp. 316-323). Multilingual Matters.
- Lavaur, J. M., & Bairstow, D. (2011). Languages on the screen: Is film comprehension related to the viewers' fluency level and to the language in the subtitles? *International Journal of Psychology*, 46(6), 455-462. <https://doi.org/10.1080/00207594.2011.565343>
- Lee, P. J., Liu, Y. T., & Tseng, W. T. (2021). One size fits all? In search of the desirable caption display for second language learners with different caption reliance in listening comprehension. *Language Teaching Research*, 25(3), 400-430. <https://doi.org/10.1177/1362168819856451>
- Li, C. (2014). An alternative to language learner dependence on L2 caption-reading input for comprehension of sitcoms in a multimedia learning environment. *Journal of Computer Assisted Learning*, 30(1), 17-29. <https://doi.org/10.1111/jcal.12019>
- Markham, P., & Peter, L. (2003). The influence of English language and Spanish language captions on foreign language listening/reading comprehension.

- Journal of Educational Technology Systems*, 31(3), 331-341. <https://doi.org/10.2190/BHUH-420B-FE23-ALA0>
- Markham, P., Peter, L., & McCarthy, T. (2001). The effects of native language vs. target language captions on foreign language students' DVD video comprehension. *Foreign Language Annals*, 34(5), 439-445. <https://doi.org/10.1111/j.1944-9720.2001.tb02083.x>
- Matielo, R., de Oliveira, R., & Baretta, L. (2017). Intralingual subtitles, interlingual subtitles, and video comprehension: Insights from an exploratory study. *Letrônica*, 10(2), 758-774. <https://doi.org/10.15448/1984-4301.2017.2.26370>
- Meara, P., & Milton, J. (2003). *X\_Lex: The Swansea levels test*. Express Publishing.
- Meara, P., & Miralpeix, I. (2006). *Y\_Lex: The Swansea advanced vocabulary levels test: v2.05*. Lognostics.
- Montero-Perez, M., Peters, E., & Desmet, P. (2014). Is less more? Effectiveness and perceived usefulness of keyword and full captioned video for L2 listening comprehension. *ReCALL*, 26(1), 21-43. <https://doi.org/10.1017/S0958344013000256>
- Montero-Perez, M., Peters, E., Clarebout, G., & Desmet, P. (2014). Effects of captioning on video comprehension and incidental vocabulary learning. *Language Learning & Technology*, 18(1), 118-141. <https://doi.org/10.125/44357>
- Muñoz, C., Pujadas, G., & Pattemore, A. (2023). Audio-visual input for learning L2 vocabulary and grammatical constructions. *Second Language Research*, 39(1), 13-37. <https://doi.org/10.1177/02676583211015797>
- Nation, P., & Heatley, A. (2002). Range: A program for the analysis of vocabulary in texts [software]. <https://www.victoria.ac.nz/lals/about/staff/paul-nation>
- Oppenheimer, J., & Arnaz, D. (1951) *I love Lucy* [TV series]. CBS.
- Peters, E. (2019). The effect of imagery and on-screen text on foreign language vocabulary learning from audiovisual input. *TESOL Quarterly*, 53(4), 1008-1032. <https://doi.org/10.1002/tesq.531>
- Pujadas, G. (2019). *Language learning through extensive TV viewing: A study with adolescent EFL learners* (Unpublished doctoral dissertation, University of Barcelona).
- Pujadas, G., & Muñoz, C. (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. <https://doi.org/10.1080/09571736.2019.1616806>
- Pujadas, G., & Muñoz, C. (2020). Examining adolescent EFL learners' TV viewing comprehension through captions and subtitles. *Studies in Second Language Acquisition*, 42(3), 551-575. <https://doi.org/10.1017/S0272263120000042>
- Pujolà, J. T. (2002). CALLing for help: Researching language learning strategies using help facilities in a web-based multimedia program. *ReCALL*, 14(2), 235-262. <https://doi.org/10.1017/S0958344002000423>



- Rodgers, M. (2013). *English language learning through viewing television: An investigation of comprehension, incidental vocabulary acquisition, lexical coverage, attitudes, and captions* (Unpublished doctoral dissertation, Victoria University of Wellington).
- Rodgers, M. (2018, October). *The effects of captions on facets of viewing comprehension of an authentic television program*. Paper presented at the Second Language Research Forum, Montreal, Canada.
- Rodgers, M., & Webb, S. (2017). The effects of captions on EFL learners' comprehension of English-language television programs. *CALICO Journal*, 34(1), 20-38. <https://doi.org/10.1558/cj.29522>
- Schmitt, N. (2010). *Researching vocabulary: A vocabulary research manual*. Springer.
- Statista. (2024). *Number of Netflix paid subscribers worldwide from 1st quarter 2013 to 1st quarter 2024*. <https://www.statista.com/statistics/250934/quarterly-number-of-netflix-streaming-subscribers-worldwide/>
- Sydorenko, T. (2010). Modality of input and vocabulary acquisition. *Language Learning and Technology*, 14(2), 50-73. <https://doi.org/10125/44214>
- Vandergrift, L. (2007). Recent developments in second and foreign language listening comprehension research. *Language Teaching*, 40(3), 191-210. <https://doi.org/10.1017/S0261444807004338>
- Vanderplank, R. (2019). "Gist watching can only take you so far": Attitudes, strategies and changes in behaviour in watching films with captions. *The Language Learning Journal*, 47(4), 407-423. <https://doi.org/10.1080/09571736.2019.1610033>
- VanPatten, B. (2002). Processing instruction: An update. *Language Learning*, 52(4), 755-803. <https://doi.org/10.1111/1467-9922.00203>
- Wang, A., & Pellicer-Sánchez, A. (2022). Examining the effectiveness of bilingual subtitles for comprehension: An eye-tracking study. *Studies in Second Language Acquisition*, 1-24. <https://doi.org/10.1017/S0272263122000493>
- Webb, S. (2010). Pre-learning low-frequency vocabulary in second language television programmes. *Language Teaching Research*, 14(4), 501-515. <https://doi.org/10.1177/1362168810375371>
- Webb, S. (2015). Extensive viewing: Language learning through watching television. In D. Nunan & J. C. Richards (Eds.), *Language learning beyond the classroom* (pp. 159-168). Routledge.
- Webb, S., & Nation, P. (2017). *How vocabulary is learned*. Oxford University Press.
- Webb, S., & Rodgers, M. (2009). Vocabulary demands of television programs. *Language Learning*, 59(2), 335-366. <https://doi.org/10.1111/j.1467-9922.2009.00509.x>
- Winke, P., Gass, S., & Sydorenko, T. (2010). The effects of captioning videos used for foreign language listening activities. *Language Learning and Technology*, 14(1), 65-86. <https://doi.org/10125/44203>

Winke, P., Gass, S., & Sydorenko, T. (2013). Factors influencing the use of captions by foreign language learners: An eye-tracking study. *Modern Language Journal*, 97(1), 254-275. <https://doi.org/10.1111/j.1540-4781.2013.01432.x>