

Geòrgia Pujadas & Carmen Muñoz (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, DOI: [10.1080/09571736.2019.1616806](https://doi.org/10.1080/09571736.2019.1616806)

Extensive viewing of captioned and subtitled TV series: a study of L2 vocabulary learning by adolescents

Geòrgia Pujadas and Carmen Muñoz

Department of Modern Languages and Literatures and English Studies, University of Barcelona, Barcelona, Spain

ABSTRACT

This study aims at exploring the potential of extensive TV viewing for L2 vocabulary learning, and the effects associated with the language of the on-screen text (L1 or L2), type of instruction (pre-teaching target items or not) and learners' proficiency. A total of 106 secondary school students (Grade 8) divided into 4 classes participated in a one-year pedagogical intervention, viewing 24 episodes of a TV series under four experimental conditions with each class being assigned to a different treatment: (1) captions and pre-teaching, (2) captions and non-pre-teaching, (3) subtitles and pre-teaching, and (4) subtitles and non-pre-teaching. Following a pre-/post-test design, form recall and meaning recall gains were examined. Results showed that participants learnt vocabulary in all four conditions, with greater gains in recalling form than in recalling form and meaning. The analysis also showed that, overall, groups that were pre-taught the target items performed better, independently of the language of the on-screen text. An important finding is the role of learners' proficiency prior to the intervention, with higher proficiency related to higher gains. The study contributes to the area of foreign language learning through audio-visual input with results from a longitudinal, classroom-based study with adolescent learners.

KEYWORDS

L2 vocabulary learning; extensive viewing; captions; subtitles; classroom instruction; adolescents

Introduction

Vocabulary is an essential part of learning a foreign language (Schmitt 2008) and

therefore a main concern for foreign language teachers. Research has shown that learners need to know around 3000 word families to understand oral discourse (e.g. van Zeeland and Schmitt 2013) and between 8000 and 9000 to understand written discourse (e.g. Nation 2006). However, in classroom settings the amount of time that can be devoted to vocabulary learning is limited, and there is a sizable gap between the amount of words that can be explicitly taught and learnt in class and those necessary to achieve higher second language (L2) proficiency (Malone 2018). Research has shown that extensive reading outside the classroom can foster vocabulary acquisition (Nation 2015; Schmitt 2008), but its impact remains limited: the typical learner does not read sufficiently to encounter the same words frequently enough to avoid forgetting them (Laufer 2005); and there has been a drop in the popularity of reading habits (European Commission 2017), especially among young people, who prefer watching TV to reading (Lindgren and Muñoz 2013; Peters 2018).

Recognising the potential limitations of reading, Webb and Rodgers (2009a) proposed extensive viewing as an alternative source of rich authentic input, pointing out the potential of television programmes and movies for learning vocabulary due to its lexical richness, repeated encounters with low-frequency words, and visual image support (Rodgers 2018; Sydorenko 2010; Webb and Rodgers 2009b). The addition of captions to this authentic input can provide access to foreign language material that would otherwise be difficult to comprehend for non-native speakers (Vander-plank 2016). Indeed, research has found that even a short amount of input enriched with captions or subtitles can lead to significant improvement in listening, content comprehension and vocabulary learning (e.g. Birulés-Muntaner and Soto-Faraco 2016; D'Ydewalle and Van Poel 1999), but there is still some debate on what language to use in on-screen text. Although typically proposed as a supplementary activity, combining explicit teaching with this media could also yield increased benefits for L2 learning, by deliberately drawing attention to specific vocabulary (Hulstijn 2013).

The emergence of multimedia learning environments and the greater accessibility to TV series, movies, and other online platforms in recent years have created opportunities for teachers and learners to boost language learning inside and outside formal settings. The aim of the present study is to explore the learning potential of audio-visual input (i.e. a TV series) with adolescent school learners.

Background

TV programmes and vocabulary learning

TV programmes have several features that make them an effective tool for L2 vocabulary learning. First of all, watching TV is already a popular activity: 88% of people in Spain watch TV daily (European Commission 2017) and if learners were to watch it in the L2 for enjoyment it could be a valuable source of meaning-focused input (Webb and Nation 2017). TV input also complies with Nation's (2007) five conditions for suitable input (Rodgers 2013): it is processed in large quantities, it is familiar to the language learners, it provides context cues (i.e. through image and dialogue), it is comprehensible (Rodgers and Webb 2011) and it is engaging. Although the association of TV programmes with entertainment has raised the concern that some learners will not pay enough attention to this type of input and will derive few benefits for language learning (Vanderplank 1990, 2016), a number of studies have shown that audio-visual material may enhance learners' motivation and attention. For instance, Pujadas and Muñoz (2017) interviewed a group of adolescent and young adult learners who had been watching an L2 TV series in the classroom. They reported that watching audio-visual materials was more 'natural', 'enjoyable' and 'motivating' than other classroom activities. This, in turn, led them to be more attentive, helped lessen anxiety and encouraged a stronger feeling of learning.

Most research on vocabulary acquisition through audio-visual input has involved incidental learning, where learning occurs as a by-product of another (meaning-focused) activity (i.e. watching a video for its information content). Indeed, a growing number of studies in this area consistently suggest that incidental vocabulary acquisition does occur through viewing short clips, full movies (Peters and Webb 2018), and TV series (Rodgers 2013). Furthermore, Rodgers and Webb (2011) found that related television programmes such as episodes in a series are likely to contain fewer word families than unrelated programmes, and that word families from the 4000 to 14000 levels were more likely to reoccur in a complete season of a television series than in a random sample of television programmes. This suggests that the more episodes you watch from the same TV series, the greater the potential to learn vocabulary from them (Webb and Rodgers 2009a). However, most studies investigating audio-visual input for L2 learning have used only short clips, segments of films, or educational videos, which are largely unrelated, and not fully representative of what a viewer might normally choose to watch (Rodgers 2013).

Only a few studies have used longer, authentic input, such as full TV episodes, documentaries or movies (e.g. Peters and Webb 2018). Longitudinal classroom-based studies using several full-length TV episodes are scarce as well. Zarei (2008) used 9 episodes of a British comedy series to assess vocabulary acquisition and comprehension.

Rodgers (2013) also investigated incidental vocabulary learning through the viewing of 10 episodes of a TV series and the effects on frequency and range of occurrence. BavaHarji, Alavi and Letchumanan (2014) used 30 episodes of a TV series to examine the effects of captioned instructional videos on EFL learners' content comprehension, vocabulary acquisition and language proficiency. Frumuselu et al. (2015) studied the acquisition of informal and conversational speech through 13 episodes of a subtitled TV series. Chen, Liu and Todd (2018) explored spoken vocabulary acquisition through 10 episodes of an animated television series. Overall – despite differences in the design and focus of these studies comparing different types of on-screen support – results indicate that incidental learning through sustained viewing does occur and that the presence of captions or subtitles is beneficial rather than distracting.

One way to optimise the effectiveness of vocabulary learning through TV programmes – and promote greater learning effort during viewing – is to involve intentional or explicit learning, on the basis of the principle that incidental and intentional learning are complementary approaches that can be integrated (Schmitt 2010). Research in the area of extensive reading suggests that learning rates can be increased by deliberately focusing attention on vocabulary (Elley 1989; Hulstijn 2013), and research on listening and the use of advanced organisers has also shown that including some kind of pre-listening support had a positive effect on comprehension (Chang and Read 2006; Chung 1996, 1999). In the area of audio-visual input, Webb (2010; see also Webb and Rodgers 2009a) investigated the potential of pre-teaching low-frequency words to increase comprehension by analysing the lexical profiles of several TV series. Webb pointed out that television programmes may be too demanding for lower level learners because they do not have the vocabulary necessary to understand the content, and suggested pre-learning unknown topic-related words in a specific television programme to improve comprehension and vocabulary learning, but to date no study has looked at the effects of teaching target expressions as part of a classroom intervention. To the authors' knowledge, the only studies including such pre-viewing activities are Bravo (2008) and Gesa and Miralpeix (2018), who investigated the learning of lexical items through several episodes of a TV series. However, since their aim was not the comparison between teaching and no teaching, no conclusions can be drawn about the effects of instruction in those studies.

In sum, while there has been an increase in studies focusing on vocabulary learning through audio-visual input, there is a scarcity of research on explicit vocabulary teaching to boost vocabulary acquisition through extensive viewing.

Captions, subtitles and proficiency

The use of original version (OV) TV programmes in the regular language classroom with children or adolescents raises the concern that learners may not be proficient enough to cope with fast speech rate and advanced vocabulary (Guillory 1998). The addition of on-screen text in the form of subtitles (native language [L1] text) or captions (L2 text)¹ – commonly available nowadays – may help these learners (Vanderplank 2016). L2 audio-visual materials enhanced with subtitles or captions are robust tools for second language learning as learners are exposed to a large amount of input simultaneously through image, text and sound. Studies on L2 learning from captioned and non-captioned audio-visual materials have consistently shown the advantage of viewing videos enhanced with on-screen text compared to viewing them without it (Mohd Jelani and Boers 2018; Montero-Pérez, Van Den Noortgate and Desmet 2013), but it is still a matter of debate whether, when, and why subtitles or captions may be preferable (Matielo, D'Ely and Baretta 2015).

The general consensus in this area is that captions provide more exposure to the target language, thus being more beneficial for language learning and vocabulary acquisition (e.g. Danan 2004; Vanderplank 2010; Winke et al. 2010). Indeed, the majority of comparative studies have found that captions have more positive effects on vocabulary learning than subtitles (Birulés-Muntaner and Soto-Faraco 2016; Frumuselu et al. 2015; Matielo, Collet and D'Ely 2013; Naghizadeh and Darabi 2015; Zarei 2008; Zarei and Rashvand 2011). Some studies do show, on the contrary, that more benefits are derived from subtitling – especially for low proficiency learners (Bianchi and Ciabattini 2008), while yet others report inconclusive results, with small or non-significant differences between captioning and subtitling (Bisson et al. 2014; Bravo 2008; Steward and Pertusa 2004).

These mixed results might be due to differences in methodology (test modality, length of exposure, target items)² and the characteristics of participants, especially their proficiency level (Malone 2018; Mohd Jelani and Boers 2018). Participants' proficiency – when reported – ranges from beginner to advanced, sometimes even within the same sample (e.g. Frumuselu et al. 2015), which poses a significant problem when discussing results against other studies (Zarei and Rashvand 2011). It has been found that learners from different proficiency levels show different responses to different on-screen text language within the same study, especially when learners are less proficient (Lwo and Lin 2012). It also appears that learners with larger vocabulary knowledge perform better than learners

with smaller vocabularies (Horst, Cobb and Meara 1998; Peter and Webb 2018; Webb and Chang 2015a), suggesting that more proficient participants will normally perform better ('the rich get richer' effect). The influence of proficiency on learners' L2 reading behaviour faced with captions and subtitles was explored in an eye-tracking study by Muñoz (2017). Unexpectedly, despite their slower reading rate, it was found that child participants with very low proficiency level spent a very short time on each fixation when words were in the L2. Muñoz suggested that learners did not even make efforts to understand because of their perceived proficiency limitations. Similarly, research on L2 reading has observed that if learners are unable to follow the overall story, they do not pay attention to the precise meaning of words (Laufer 2005). If a minimum competency threshold is necessary to benefit from captioning (Neuman and Koskinen 1992), subtitling might be a better option for supporting beginner learners (Danan 2004). However, the use of subtitles is normally discouraged in foreign language classroom settings, because it is believed that the availability of the L1 will stop learners from listening to the foreign language and that they will focus only on reading (Danan 2004). It has been argued, though, that 'it seems perfectly sensible to exploit it [the L1] when it is to our advantage' (Schmitt 2008:337), which seems particularly true for younger viewers.

Young viewers

Studies on audio-visual input and L2 learning have been frequently conducted at university level or with adult language learners (e.g. Montero-Perez et al. 2014; Sydorenko 2010; Winke, Gass and Sydorenko 2010). Although still scarce, the number of studies focusing on multimodal input with children and adolescents has increased in the last two decades, and research has demonstrated that watching subtitled or captioned television has positive effects on both first and foreign language learning. Early studies observed that primary school children benefitted from subtitles in their L1 (Koolstra and Beentjes 1999), and that even pre-schoolers could learn new L1 vocabulary through exposure to audio-visual input (Rice et al. 1990). D'Ydewalle and Van Poel (1999) conducted a study with 8- to 12-year-olds testing incidental L2 learning through a 10-minute still-motion movie comparing normal and reversed subtitles, and found that – even with a short exposure – participants in both conditions already showed small gains in vocabulary. From data collected through questionnaires, Kuppens (2010) also reported that Grade 6 students who frequently watched subtitled English programmes (before formal instruction) performed significantly better in English tests.

The influence of frequent watching of audio-visual material over other types of out-of-

school exposure has been observed in several studies as well. Lindgren and Muñoz (2013) found that watching movies had the strongest explanatory power on the listening and reading comprehension skills of a very large group of 10- to 11-year-old learners in seven European countries. Similarly, in a study comparing Danish and Spanish learners of English (ages 7 and 9) Muñoz, Cadierno and Casas (2018) found that Danish children's more frequent viewing of OV audio-visual material helped compensate for the effects of their comparatively lower amount of formal instruction to a greater extent than other activities such as gaming or listening to music. Peters (2018) showed that 40% of surveyed students (15- to 16-year-olds) watched OV TV several times a week, compared to the 1% who indicated reading books with the same frequency.

Several recent studies have focused on the comparison between young learners' viewing audio-visual material with or without subtitles or captions. Hsu et al. (2013) investigated the effect of subtitle mode on vocabulary and comprehension with Grade 5 participants during a one-month experiment, comparing non-captioning, full-captioning and keyword-captioning. These researchers found that there were no differences between the two captioning groups and that both outperformed the non-captioning one. Lekkai (2014) explored incidental vocabulary learning through a 15-minute cartoon with Grade 4, 5 and 6 learners (ages 9–12), with and without L1 subtitles. Again, learners in the subtitles group outperformed the non-subtitles and control groups, supporting the idea that even at this young age students can learn from subtitled videos. Chen, Liu and Todd (2018) explored the effect of captioning (against non-captioning) on spoken vocabulary with Grade 8 learners (aged 13), and found that the availability of captions significantly improved learners' recognition of form and form-meaning knowledge of novel L2 spoken vocabulary, especially for higher-proficiency learners.

Studies comparing the effects of on-screen text language (either L1 or L2) have also been recently conducted. Bravo (2008) compared the effects of watching captioned or subtitled episodes of a TV show on lexical expressions and comprehension scores for 13- to 14-year-olds. She found similar results for both experimental groups but also reported that the absence of the L1 required a greater effort and higher L2 fluency among her young participants when completing post-viewing tasks. Lwo and Lin (2012) compared L1 and L2 text, using a multimedia animated reading tool to explore the effects of different types of on-screen text (L1, L2, L1 + L2 and none) on vocabulary and reading comprehension with Grade 8 learners. They found that the effects of different modes on scores depended on learners' L2 proficiency, that for the lower-proficiency learners having L2 or L1 + L2 subtitles was more beneficial, and that learners relied on visual information for comprehension. Naghizadeh and Darabi (2015)'s study on L2 vocabulary with

intermediate-level 15- to 17-year-olds reported that learners in the captions groups learnt significantly more than those in the subtitles group, who in turn had similar results to the non-subtitles group. Peters, Heynen and Puimège (2016) conducted two experimental studies on the effects of L1 and L2 text on vocabulary gains for 17- to 18-year-olds (intermediate and low-intermediate). Their results showed that, even if gains were low, captions had the potential to increase form learning, and that the captions group outperformed the subtitles group.

Altogether, the above studies suggest that, regardless of subtitling mode, length of exposure to the input, or proficiency, young learners benefit from exposure to audio-visual input enhanced with L1 and L2 on-screen text. As with older learners, results seem to indicate that captions are more suitable for older/more proficient young learners, while subtitles might be more appropriate for younger/less proficient children.

Aims and research questions

The present study aims to address some of the gaps observed in this field of research, as highlighted above. It focuses on L2 vocabulary learning by adolescents through extensive viewing of an OV TV series over a complete academic year, providing much needed longitudinal data. It compares vocabulary learning gains from the same audio-visual material with two types of text support: captions and subtitles, which will allow comparison of the benefits of the two modes for beginner / low-intermediate learners and exploration of the role of learners' proficiency in this learning environment. The study also aims to compare vocabulary learning gains under two different instructional conditions: focused (with pre-teaching of target items) and non-focused (without pre-teaching),³ thus contributing to the understanding of the pedagogical value of both types of instruction. Specifically, the study addresses the following research questions:

(1) To what extent can L2 vocabulary (form and meaning) be learnt through extended exposure to TV series?

(2) How is L2 vocabulary learning through TV series affected by (a) language of the on-screen text, (b) type of instruction and (c) learners' proficiency level?

Methodology

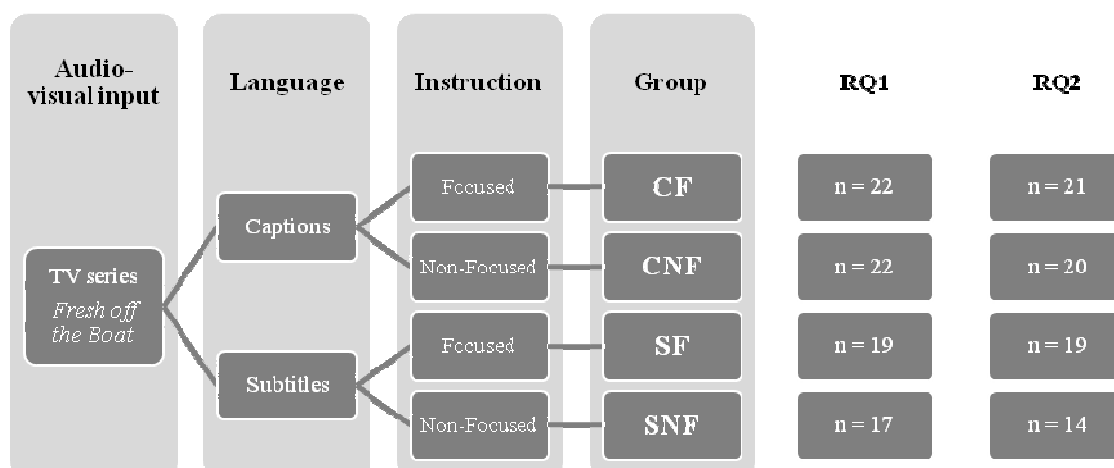
Participants

The initial pool of participants comprised 106 secondary school learners (65 female, 41 male) in Grade 8, from a state school in the Barcelona area. Only participants who had 85% attendance or more per term were included in the analysis (N = 80). For the second research question, participants who had not completed the proficiency test had to be excluded, leaving a total of 74 (46 female, 28 male). They were 13–14 years old at the time the intervention started.

Participants were Catalan-Spanish bilinguals, most of them balanced bilinguals for whom both languages may be considered first languages. They had a beginner – low-intermediate proficiency level in English (Pre-A to A2/B1 according to the Common European Framework of Reference [CEFR]) and a mean vocabulary size in English of 1,967 words. Prior to the intervention, around 55% of participants reported watching movies or TV series in English with L1 subtitles on a weekly basis, and around 15% with L2 subtitles or no subtitles. More than 50% said they found subtitles to be useful or very useful while only 4% considered them to be useless or annoying.

The participants had been randomly distributed in four classes by the school, and each one of them was assigned to a different experimental condition (see Figure 1): captions and focused instruction (CF) (n = 22); captions and non-focused instruction (CNF) (n = 22); subtitles and focused instruction (SF) (n = 19); and subtitles and non-focused instruction (SNF) (n = 17). There were no significant differences between the four groups either in proficiency or vocabulary size (see below).

Figure 1. Experimental conditions



CF = captions focused; CNF = captions non-focused; SF = subtitles focused; SNF = subtitles non-focused

Audio-visual material

The TV series selected for the intervention was *Fresh off the Boat* (Khan et al. 2015), and it was chosen for the following reasons: the episodes had an appropriate length (with an average running time of 20 minutes); its content was appropriate for this particular age group; it was engaging (participants could identify with the main character and get hooked into the story); it was serial in nature, which allowed participants to gather information about the characters as they continued watching new episodes (Rodgers 2013); it was not strongly accented; it was a sitcom, a format with which participants are familiar through watching similar TV programmes; and at the time the intervention took place *Fresh off the Boat* had not been aired in Spain, thus minimising the possibility that participants had watched any of the episodes before.

From the series, 24 consecutive episodes were selected, and by the end of the intervention participants had watched a total of 8 hours and 35 minutes of audio-visual input. Subtitles (in Spanish) were manipulated by the first author to ensure that the number of encounters with the target items was equal in both captions and subtitles conditions. The 24 episodes chosen for the intervention were analysed using the RANGE software (Nation and Heatley 2002). The analysis of the lexical profile shows that the series reached a 95.66% coverage at the 3000-word level plus proper nouns and marginal words.⁴

Target items

A total of 120 target items (TIs) (40 each term, 5 per episode) were selected from the series according to (a) frequency of occurrence (between 2 and 14 times within the episode) and (b) low likelihood of being known by participants at this level of proficiency (school teachers were consulted on the TIs selected and these were replaced when necessary).⁵

TIs were from the first to the fourteenth frequency word lists on the BNC/COCA (Nation 2012): 52% belonged to the 1–3 K word families, 21% to the 4–8 K, and 12% to the 9–19 K (15% were off-list). TIs also belonged to different parts of speech, with the majority of them being nouns (60%) and verbs (25%). As for the frequency of occurrence across the intervention, 75% occurred between 2 and 5 times, 20% between 6 and 9 times and only 5% between 10 and 14 times.

Test instruments

Initial general proficiency was assessed by means of the Oxford Placement Test (OPT) and vocabulary size was measured using X_Lex (Meara and Milton 2003). Two questionnaires were also administered prior to the intervention to collect background information and attitudes towards subtitles (see above). A third questionnaire was administered after the intervention to gather participants' insights about the perceived usefulness and overall learning value of the viewing sessions.

The pre- and post-test assessing learners' knowledge of the TIs consisted of two parts: (1) an aural form recognition and written form recall test (henceforth form recall), and (2) a meaning recall test (henceforth meaning recall). Participants listened to each TI twice and had to write down the English word, and then provide a translation or a short definition in Catalan or Spanish. Tests were administered by the first author and included a set of trial items to ensure participants completed them correctly. This type of test was chosen because, in order to assess the benefits of captions and subtitles for vocabulary learning, tests had to be congruent with the input-modality (Mohd Jelani and Boers 2018): written L2 word prompts in the test could have been used to the caption groups' advantage.

Procedure

The classroom intervention took place over a whole academic year and was embedded as a part of normal English lessons. Because of school calendar constraints, the intervention itself was divided into 3 terms with 8 viewing sessions each, and participants were pre-tested at the beginning and at the end of each term to assess their knowledge of the corresponding 40 TIs (Figure 2 shows the structure of the intervention). Pre-tests were administered 1–2 weeks prior to the first session to reduce pre-test effects. The decision to have three sets of pre-/post-tests was also made to avoid decay due to having the post-test too far from the first sessions of the term. The complete inter- vention extended over 32 sessions.

Figure 2. Structure of the intervention

Prior to intervention	TRIMESTRE 1 September – December			TRIMESTRE 2 January – March			TRIMESTRE 3 April – June		
	Proficiency tests	Pre-test	TREATMENT	Post-test	Pre-test	TREATMENT	Post-test	Pre-test	TREATMENT

Questionnaires								
----------------	--	--	--	--	--	--	--	--

For the two focused instruction groups (one with captions, one with subtitles), each viewing session started with a pre-viewing task aimed at teaching the five TIs appearing in the episode plus three distractors. Pre-viewing activities included matching exercises, word searches, fill-in-the-blanks tasks and crosswords, and they were corrected orally. These activities had the aim of drawing learners' attention to the TIs while watching the video, but no specific strategies were suggested. Then, participants watched the episode (with either captions or subtitles, depending on the group) and completed two immediate post-viewing tasks, namely a vocabulary task and a content comprehension task,⁶ which were given to encourage learners to pay attention to both vocabulary and content. The two non-focused instruction groups' sessions followed the same outline but did not include the pre-viewing task; thus, learners were unaware of what words they were going to be tested on later.

Scoring and analysis

Pre- and post-tests were scored dichotomously (1 or 0). For form recall, words had to be correctly spelled to be considered correct. For meaning recall, translations were scored by two raters (there was an interrater reliability of 95%; disagreement cases were discussed until an agreement was reached). A word was considered learned when it was unknown in the pre-test and known in the post-test. Words known in both pre- and post-test were considered known but not learned. Relative gains were calculated at item level following the formula used in previous studies (Horst, Cobb and Meara 1988; Peters and Webb 2018; Rodgers 2013):

$$\text{Relative gains} = (\text{number of learnt TIs} / \text{total number of TIs} - \text{number of known TIs}) \times 100$$

Prior to conducting the analysis, as stated above, participants with less than 85% attendance at the viewing sessions across the intervention were excluded from the data. TIs corresponding to missing episodes were not taken into account when calculating the vocabulary gains for those participants. The measure of relative gains used for analysis is the average relative gains across three terms.

Results

Table 1 shows descriptive statistics for the two proficiency measures: OPT and X_Lex. There were no significant differences between groups in terms of proficiency ($F(3,70) = 1.545, p = .210$) nor vocabulary size ($F(3,64) = .816; p = .490$). For the present study, OPT scores (henceforth proficiency) were used because the OPT yields a more general measure of proficiency, including a section on listening, which was deemed appropriate considering that listening skills are especially relevant in this learning environment.

Table 1. General proficiency and Vocabulary size descriptives

Group	General proficiency		Vocabulary size	
	n	OPT scores	n	X_Lex score
CF	21	99.71 (14.05)	17	1971 (547)
SF	19	90.47 (13.07)	16	2097 (332)
CNF	20	92.75 (15.04)	19	1992 (601)
SNF	14	93.43 (15.33)	16	1825 (434)
Average		94.27 (14.49)		1972 (494)

Research question 1: To what extent can vocabulary be acquired through TV series?

Our first research question focused on the extent to which participants could learn L2 vocabulary through an extended exposure to TV series shown in the classroom. The descriptive statistics for relative gains in form recall and meaning recall are presented in Table 2. The table displays the percentage of relative gains (average across the three terms) – with standard deviations shown in brackets – by experimental group, and the average across the four groups.

Table 2. Relative gains (in percentage) for form recall and meaning recall

Group	n	Percentage Relative gains (SD)	
		Form	Meaning
CF	21	30,10 (16,83)	14,54 (10,43)
SF	19	21,53 (11,16)	8,45 (6,36)

CNF	20	13,02 (5,85)	5,97 (6,25)
SNF	14	14,30 (9,32)	8,34 (7,73)
Average		20,29 (13,50)	9,49 (8,48)

Overall, the focused groups who were pre-taught the TIs performed better than the non-focused groups, with the captions-focused group (CF) being the most successful, followed by the subtitles- focused group (SF), the subtitles non-focused group (SNF) and finally the captions non-focused group (CNF).

As can be observed, for form recall the two non-focused groups – who were not told what words to pay attention to – performed similarly, independently of the language of the on-screen text. In contrast, when learners were pre-taught the words, the CF group – with simultaneous exposure to both L2 sound and text – outperformed the SF group. For meaning recall, however, the two subtitles groups – with access to L1 translations – performed similarly, independently of instruction.

Paired-sample t-tests showed that differences between pre- and post-test were significant in all three terms for both form recall and meaning recall ($p < .001$) and for each experimental group (ranging from $p < .001$ to $p = .029$). On average, the most successful group had gains in form recall of 30.10% and gains in meaning recall of 14.64%, which means that on average participants in that group learnt approximately 36 word forms and 18 word meanings. It is interesting to note that the differences in gains within the group are considerable, with the most successful participants having gains of 62.18% in form and 32.11% in meaning, while the least successful ones had gains of 5.4% in form and 0.83% in meaning.

Differences between the groups in form recall were first explored by means of a Welch's ANOVA and using squared transformations (a Levene's test showed that variances were unequal: $F(3,76) = 2.774$, $p = .047$). The ANOVA showed that there was a statistically significant difference between groups ($F(3,76) = 7.714$, $p < .001$, $\omega^2 = .199$) and that approximately 20% of the total variance in form recall gains was accounted for by the experimental group. A Tamhane's T2 post-hoc test revealed that there was a statistically significant difference in relative gains in form between CF and CNF groups ($p = .001$) as well as between CF and SNF groups ($p = .003$). For meaning recall, a one-way ANOVA showed that there were also significant differences between the experimental groups ($F(3,76) = 3.301$, $p = .024$). A Tukey HSD post-hoc test revealed that differences were only significant between CF and CNF groups ($p = .023$).

Research question 2: Effect of language, type of instruction and proficiency

Our second research question addressed the extent to which language of the on-screen text, type of instruction and learners' proficiency could explain the differences in vocabulary learning observed among the four groups. Generalised linear models (GLMs) were run to evaluate the influence of these three factors on the two vocabulary outcome measures: form recall and meaning recall. Exploration of the data showed that the variable proficiency was not linearly distributed. For this analysis, proficiency was re-categorised into three levels according to the OPT scores, distributing participants in three CEFR groups: Pre-A (n = 25), A1 (n = 35) and A2/B1 (n = 14).

A GLM was first calculated with form gains (relative gains in percentage for form recall across the intervention) as the dependent variable, and language (captions or subtitles) type of instruction (focused or non-focused), and proficiency (Pre-A, A1 or A2/B1) as fixed effects. Non-significant inter- actions were removed from the model, leaving significant main effects. [Table 3](#) presents the fitted model for relative gains in form.

Table 3. Results of fitted generalized linear model: influence of fixed factors on form recall

	Mean (SE)	M Diff (SE)	df	F	Sig.
Captions (EN)	22.25 (1.75)	2.53 (2.52)	1, 66	1.012	.318
Subtitles (SP)	19.72 (1.95)				
Focused	26.93 (1.78)	11.89 (2.50)	1, 66	22.596	.000
Non-Focused	15.04 (1.91)				
Pre-A ^a	14.65 (2.13)	^{a-b} 4.30 (2.79)			
A1 ^b	18.95 (1.81)	^{b-c} 10.41 (3.45)	2, 66	8.164	.001
A2/B1 ^c	29.35 (2.96)	^{a-c} 14.70 (3.65)			

The model revealed a main effect for type of instruction and proficiency, but no main effect for the language of the on-screen text. Results showed that participants in the focused condition scored significantly higher ($p < .001$) than their classmates by about 11.89%. Gains in form also depended significantly on participants' proficiency level ($p=.001$), with the more proficient students scoring 14.70% higher than the less proficient students.

Once it was established that these two factors had a significant effect on learning outcomes, we further explored whether the effect of learners' proficiency was different in

each instruction group. For the following analysis, the CF and SF groups were jointly considered focused, and CNF and SNF were non-focused (for we did not find a significant effect for on-screen language). Table 4 shows the percentage of relative gains per type of instruction and proficiency group, with standard error (SE) in brackets.

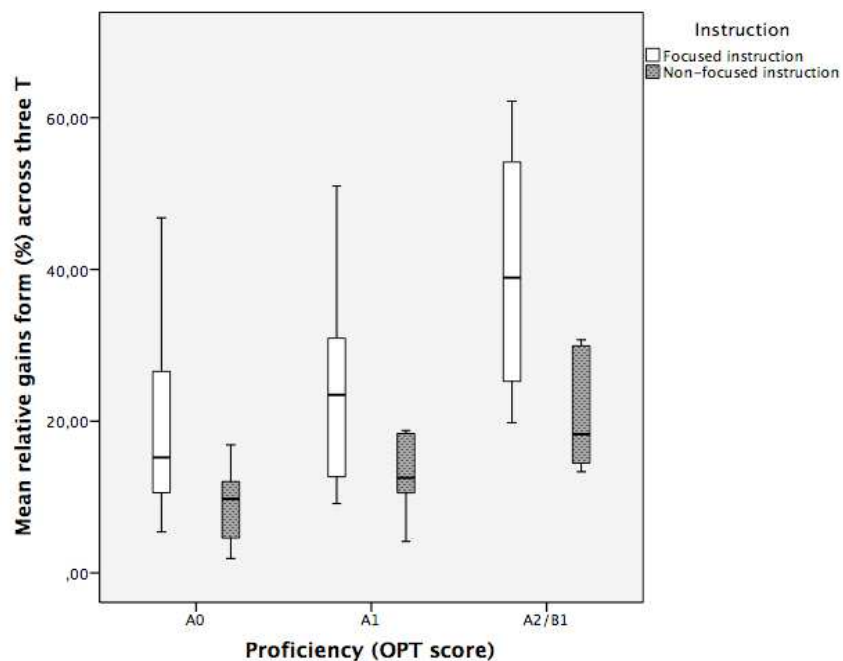
Table 4. Relative gains in form and meaning per type of instruction and proficiency group

Type of Instruction	Proficiency	n	Percentage of relative gains Mean (SE)	
			Form	Meaning
Focused	Pre-A	13	20.44 (2.99)	8.08 (2.01)
	A1	19	24.17 (2.46)	10.61 (1.65)
	A2/B1	8	37.57 (4.09)	18.89 (2.75)
Non-focused	Pre-A	12	8.86 (3.10)	2.73 (2.09)
	A1	16	13.90 (2.70)	6.99 (1.82)
	A2/B1	6	20.83 (4.37)	14.65 (2.95)

GLM results showed that there were significant differences between types of instruction when comparing each proficiency level against its counterpart, with the focused group significantly outperforming the non-focused group at the Pre-A ($F(1,66) = 7.182, p = .009$), A1 ($F(1,66) = 7.928, p = .006$) and A2/B1 ($F(1,66) = 7.820, p = .007$) levels.

Within the focused group itself, results indicated that differences were also significant between the three levels of proficiency ($F(2,66) = 5.880, p = .004$), and pairwise comparisons showed that significant differences were found between the A2/B1 level and both Pre-A ($p = .001$) and A1 levels ($p = .006$). In contrast, in the non-focused group differences were marginally significant ($F(2,66) = 2.553, p = .087$) and significant differences were found only between the A2/B1 level and the Pre-A level ($p = .029$). Figure 3 shows estimated marginal means per focused and non-focused groups when we divided participants by proficiency levels (Pre-A, A1 and A2/B1).

Figure 3. Estimated marginal means for form recall of focused and non-focused groups



A second GLM was calculated with meaning gains (relative gains in percentage for meaning recall across the intervention) as the dependent variable, and language (captions or subtitles), type of instruction (focused or non-focused) and proficiency (Pre-A, A1 or A2/B1) as fixed effects. Again, non-significant interactions were removed, leaving significant main effects in the model. Table 5 presents the fitted model.

Table 5. Results of fitted generalized linear model: influence of fixed factors on meaning recall

Mean (SE)	M Diff (SE)	df	F	Sig.
-----------	-------------	----	---	------

Captions (EN)	10.91 (1.75)	1.18 (1.69)	1, 66	.488	.487
Subtitles (SP)	9.73 (1.31)				
Focused	12.48 (1.19)	4.32 (1.68)	1, 66	6.643	.012
Non-Focused	8.16 (1.28)				
Pre-A ^a	5.43 (1.42)	^{a-b} 3.35 (1.87)			
A1 ^b	8.77 (1.21)	^{b-c} 8.00 (2.31)	2, 66	10.810	.000
A2/B1 ^c	16.77 (1.99)	^{a-c} 11.34 (2.45)			

Similar to gains in form, the model showed a main effect for type of instruction and proficiency, but no main effect for the language of the on-screen text. Results showed that participants in the focused condition scored significantly higher ($p=.020$) than their classmates by 4.32%. Gains in meaning also depended significantly on participants' proficiency level ($p < .000$), with the more proficient students scoring 11.34% higher than the less proficient students.

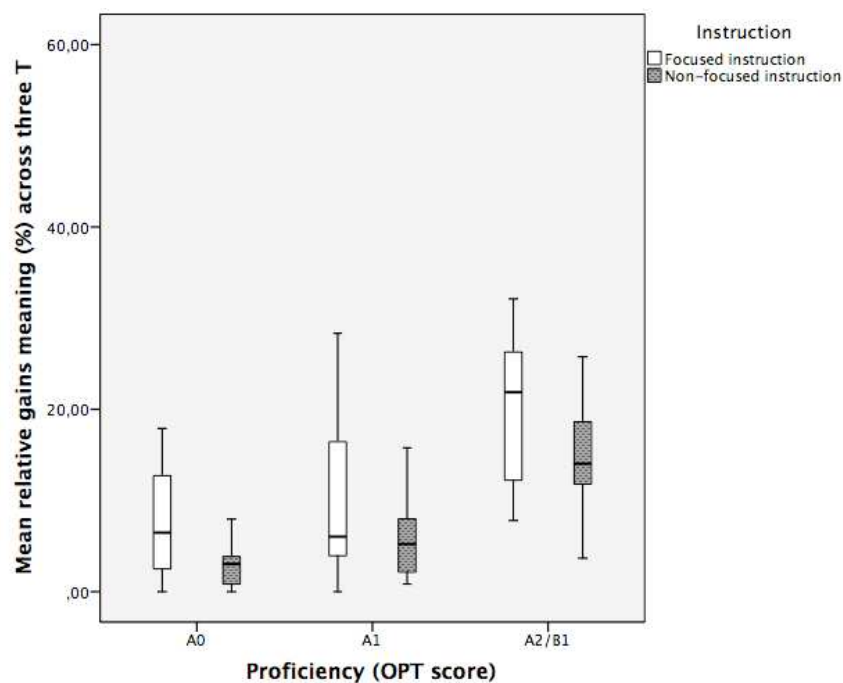
Once more we examined the relationship between instruction and proficiency (see [Table 4](#)). In contrast with form recall, for meaning recall there were no significant differences between instruction groups at the Pre-A ($F(1,66) = 3.375, p = .071$), the A1 ($F(1,66) = 2.179, p = .145$) nor the A2/B1 level ($F(1,66) = 1.103, p = .298$), although learners in the focused group always had higher gains than those in the non-focused group.

GLM also showed, however, that differences in proficiency were significant within each instruction group ($F(2,66) = 5.111, p = .009$ for focused, and $F(2,66) = 5.452, p = .006$ for non-focused). Pairwise comparisons revealed that in both conditions it was the more advanced group (A2/B1 level) that significantly outperformed the other two: in the focused group, differences were between the A2/B1 level and the Pre-A ($p = .002$) and A1 levels ($p = .012$), and in the non-focused group, significant differences were also found between the A2/B1 group and the Pre-A ($p = .002$) and A1 ($p = .006$) groups. [Figure 4](#) shows estimated marginal means per type of instruction when we divided participants by proficiency levels.

Finally, and although a complete analysis falls beyond the scope of this paper, a preliminary exploration of the self-reported data from the end-of-intervention questionnaires revealed interesting findings concerning learners' attention to the audio-visual input. As many as 89% of participants reported having been attentive or very attentive, and only 17% said they were paying less attention during the viewing sessions

compared to other classroom activities. Learners also reported that they understood better the series by the end of the intervention (74%), they had a strong feeling of learning (58%) and they felt relaxed during the sessions (53%) (Pujadas, 2019).

Figure 4. Estimated marginal means for meaning recall of focused and non-focused groups



Discussion

This study explored the effects of extensive viewing of a TV series on L2 vocabulary learning investigating the influence of the language of the on-screen text, type of instruction and learners' proficiency on L2 vocabulary learning. Results concerning the first research question showed that, independently of the experimental condition,

participants did learn L2 vocabulary from extensive exposure to audio-visual input. This concurs with findings from the majority of studies in the field, which claim that L2 vocabulary can be acquired through watching TV series (e.g. Peters and Webb 2018; Rogers 2013). Additionally, it was also found that participants made greater gains in recalling form than in recalling meaning – in line with previous studies that also showed greater gains in form recognition than in meaning recognition (Montero-Perez et al. 2014; Peters, Heynen and Puimège 2016). This was found across all conditions, that is, with captions or subtitles, and with or without pre-teaching of TIs. Importantly, it cannot be forgotten that in the present study, recalling meaning is dependent on prior recalling form, which will have had an effect on meaning recall scores.

Our second research question looked at the effect that the language of the on-screen text, type of instruction and learners' proficiency had on participants' vocabulary learning regarding form recall and meaning recall. Overall, focused instruction groups performed significantly better than non-focused groups in both form recall and meaning recall, independently of whether they were watching the series with captions or subtitles. This is not surprising, since it is well known from past research that intentional learning is significantly more efficient than incidental learning (Hulstijn 2003).

However, although results revealed that language of the on-screen text had no significant effect on either form recall or meaning recall, a pattern can be observed: in the focused condition, the group with captions outperformed the group with subtitles in both form and meaning recall, while in the non-focused groups it was the subtitles group that performed slightly better than the captions group in meaning recall. This may suggest that when learners are pre-taught the words appearing in an episode they make a first connection between form and meaning of the new words through the pre-viewing activities. Then, having the audio and text in the same language (captions) reinforces the connection between the oral and written form (Webb and Nation 2017), which in turn helps recall the meaning. On the other hand, when learners are not pre-taught the TIs (more comparable to an incidental learning condition), there is a tendency for the subtitles group to have higher gains in meaning, as they can use the meaning provided by the L1 subtitles to connect it to the L2 oral form, but cannot use this shortcut with captions (and it takes them longer to learn the words). This would also suggest that the L1 text might have compensated for the lack of instruction in the SNF group.

The lack of statistical differences between the captions and subtitles groups falls in line with results from other aforementioned studies (e.g. Bisson et al. 2014; Steward and Pertusa 2004). If we narrow down the comparison to studies with young viewers, results

coincide with those from Bravo's study (2008) – with participants at A2/B1 proficiency level – in which L1 and L2 groups did not statistically differ, though the subtitles group performed slightly better. However, Bravo acknowledged that the L1 group was initially more proficient than the L2 group, and since the presence of captions required a higher L2 proficiency level this would explain the lack of differences. Lwo and Lin (2014) also found that varying the language of the text did not have a significant impact on vocabulary gains and that the effect of different types of text presentation varied depending on learners' proficiency. They found that this was more evident in lower proficient learners, who benefited the most from captions, but for advanced learners, the presence of the L1 was a distractor – a result that was not found in the present study.

On the other hand, the studies by Naghizadeh and Darabi (2015) and by Peters, Heynen and Puimège (2016) consistently found that captions groups performed significantly better than the sub- titles groups in vocabulary learning. However, learners in these two studies were older (aged 15–18) than participants in the present study (aged 13), and more proficient. This suggests that there might indeed be an age/proficiency threshold and that the older and more proficient you are, the better you benefit from captioning rather than subtitling.

In line with the above-mentioned studies, in our study, we found that – as expected – learners' proficiency level was significantly related to vocabulary gains in both form and meaning recall, with more advanced learners obtaining higher gains (Chen, Liu and Todd 2018). Since instruction had a strong effect on vocabulary learning outcomes, we further investigated the relationship between learners' proficiency and instruction and found that the effect of proficiency level (Pre-A, A1 or A2/B1) was different depending on whether TIs were pre-taught or not. Results showed that for form recall, participants in the focused groups significantly outperformed the non-focused groups at each proficiency level and that the A2/B1 group had higher significant gains than A1 and Pre-A. In the non-focused groups, significant differences were only found between A2/B1 and Pre-A levels.

In contrast, for meaning recall, the differences between focused and non-focused groups at each proficiency level did not reach significance, although focused groups consistently outperformed non- focused groups. This would suggest that for meaning recall, proficiency might have had a slightly stronger effect than instruction. Again, A2/B1 groups in both types of instruction setting significantly outperformed A1 and Pre-A groups. The fact that significant differences were mostly found between the A2/B1 group and the other two suggests the possibility of a threshold at A2/B1 level, over which

learners seem to be able to benefit better from exposure to audio-visual input for vocabulary learning. This finding confirms the crucial role played by proficiency and suggests that the results of studies in this area cannot be adequately interpreted if this key variable is not taken into account. In other words, to reach robust conclusions in this line of research, results need to be seen as contingent on the proficiency level of the participants of each particular study. Further research controlling for age or for proficiency can help us conclude which of the two factors has a stronger interaction with the outcomes from either mode.

In summary, this study has yielded evidence that shows that extensive viewing of TV programmes can support L2 vocabulary learning. Arguably, compared to the outcomes of other kinds of vocabulary instruction, the gains in both form and meaning were relatively small (in a total of 515 minutes, participants in the most successful group (CF) learned on average around 36 word forms and 18 word meanings). However, one needs to bear in mind that vocabulary learning is a gradual process and we did not test the partial acquisition of the TIs nor possible long-term effects. Likewise, we did not test for other aspects of knowledge, such as pronunciation, or other words that might have been learned.

Conclusions

This study has made several contributions to the growing area of research into the benefits of watching captioned/subtitled audio-visual material for L2 learning. First, results confirm previous findings regarding the potential of TV programmes for vocabulary learning and as a rich source of comprehensible input. Secondly, this study has proven that the integration of explicit instruction and extensive viewing is possible and effective, and suggests that a small amount of teaching (instruction consisted of simple 5 minute-activities) and directing learners' attention to target vocabulary already brings about significant improvement – especially on form recall. Third, it has provided valuable data for the so-far unresolved issue of the relative gains from captions and subtitles, underscoring the key role played by learners' proficiency. In that respect, it has been found that the benefits of either on-screen text language depend not only on the linguistic competence itself but also on the instruction condition. Fourth, the study looked at adolescent learners in regular classrooms and measured gains over a period of time extending a whole academic year, yielding ecological validity and generalisability to its findings and complementing results more frequently obtained from research with university students, although replication with larger groups is still needed.

Finally, relevant pedagogical implications emerge from this classroom-based intervention. As mentioned above, a minimal time investment on pre-teaching with receptive activities helps students' learning of vocabulary, which could be further enhanced with additional productive activities (Chung 1996; Sockett 2014). Data on learners' perceptions also showed that learners were very positive and appreciative of encountering 'real' English, which is of special interest for teachers. Students were paying attention to the input and had a strong feeling of learning, and the entertaining nature of the media did not seem to distract but, on the contrary, boosted attention. Moreover, although results evince the higher efficiency of pre-teaching, the fact that learners who did not receive instruction also succeeded in learning vocabulary – even if to a lesser extent – supports previous findings on the learning potential that watching TV series may also have outside the classroom. Especially in settings where L2 input is limited, teachers may encourage learners to engage in viewing OV series out of school and may also train them to enhance their learning potential through the use of strategies and focus on form. Finally, the study also has social implications demonstrating to learners and teachers the value of watching captioned/subtitled TV series as an L2 resource in a traditionally dubbing country and, thus, hopefully contributing to a change of viewing habits.

The study has some limitations that should be acknowledged. First, the type of test used to evaluate learning might explain the low gains obtained – especially in meaning – since a recall test (e.g. a translation test) is more difficult than a recognition test (e.g. multiple-choice test) (Jones 2004). If a student failed to aurally identify the target items in the first place they could not provide a translation, but this does not mean the participants could not recognise the word form if encountered, or that they did not know the meaning of the word. A second test to check partial knowledge of the target items may have provided a more accurate picture of their learning. As regards form, the requirement that words had to be correctly spelt might have put the subtitles groups in disadvantage. Another limitation is that the study did not take into account the degree of attention participants were actually putting into the tasks, which could be a concern when extrapolating results. Although more faithful to the real learning environment, this is a common short-coming of classroom studies, which may be seen as an inevitable concomitant of their ecological validity. Similarly, this environment precluded the existence of a control group as ethically, their learning opportunities without any exposure to OV TV might have been reduced. Lastly, we did not take into account the role that the imagery might have had in providing semantic support, especially for those learners who did not receive that information from either the instruction or the subtitles (the CNF group), and who nonetheless acquired a number of the word meanings. Future research also needs to look into the characteristics of the target items – such as imageability (Rodgers 2018),

frequency of occurrence or relevance, which may influence the learnability of vocabulary in this particular context.

Notes

1. In the present study, the term subtitles is used to refer to L1 subtitles or interlingual subtitles (in this case, Spanish), and captions is used to refer to same-language subtitles or intralingual subtitles (in this case, English).

2. Within the aforementioned studies input materials vary from 2-minute short clips to full-length 90-minute movies or several TV episodes (adding up to 325 minutes in total), and the number of target items oscillate between 10 and 78 items, with a wide range of tests types and constructs measured.

3. The terms focused and non-focused are used here to refer to the condition in which learners are pre-taught the target items and are not, respectively. These terms are preferred over the terms intentional and incidental. Usually, incidental learning is narrowly identified as not having forewarning of an upcoming post-test (Dörnyei 2009; Hulstijn 2003), and the distinction with intentional learning is centred on the role of learners' intention (Burton, García López and Esquileche Mesa 2011). However, in classroom studies 'it is difficult [...] to ensure that learners do not become intentionally focused on learning vocabulary' (Malone 2018, 3), and it is actually impossible to know what participants actually do (Hummel 2010). This becomes even more obvious in the context of a longitudinal study, where participants are already expecting a vocabulary test after a couple of sessions.

4. Nation (2006) suggested that proper nouns have a minimal learning burden; and Webb and Rodgers (2009b) showed in their study of the lexical coverage of movies that if learners knew proper nouns and marginal words (e.g. ah, oh, huh) they could reach 95.76% coverage with the most frequent 3,000 word families. In the present study, proper nouns make up 3.11 % of the running words, adding more coverage than words from the 3000 word-family (1.62%). Considering that characters and locations reoccur throughout the episodes, it seems safe to assume learners are familiar with most of the proper nouns.

5. Frequency of occurrence across the intervention takes into account the number of occurrences of a TI from the pre- to the post-test where the TI was tested, that is, it includes occurrences happening within 8 episodes of the term in which that word was targeted.

6. The vocabulary task was an aural recall and meaning recognition task (participants had to write down

the word they heard, and select the correct translation from 5 options provided), and included 5 TIs plus 3 distractors. The comprehension task included 10 items (5 multiple-choice, 5 true-false) assessing content comprehension. Post-viewing tasks were not corrected. Although the analysis of comprehension falls beyond the scope of the present paper, it is interesting to note that, while comprehension was observed in all groups, differences were found between on-screen text language, with the subtitles groups outperforming the captions groups (Pujadas and Muñoz, in press).

Disclosure statement

No potential conflict of interest was reported by the authors.

Funding

This work was supported by Ministerio de Economía y Competitividad (MINECO) [grant number FFI2016-80564-R] and by Agència de Gestió d'Ajuts Universitaris i de Recerca (AGAUR) [grant numbers 2017 SGR560 and FI-DGR 2016].

ORCID

Geòrgia Pujadas <http://orcid.org/0000-0002-0290-1158>

Carmen Muñoz <http://orcid.org/0000-0002-7001-4155>

References

- Bavaharji, M., Z.K. Alavi and K. Letchumanan. 2014. Captioned instructional video: effects on content comprehension, vocabulary acquisition and language proficiency. *English Language Teaching* 7, no. 5: 1–16. DOI: 10.5539/elt.v7n5p1.
- Bianchi, F. and T. Ciabattini. 2008. Captions and subtitles in EFL learning: an investigative study in a comprehensive computer environment. In *From Didactas to Ecolingua*, ed. A. Baldry, M. Pavesi, C. Taylor Torsello and C. Taylor, 69–90. Trieste: Edizioni Università di Trieste.
- Birulés-Muntané, J. and S. Soto-Faraco. 2016. Watching subtitled films can help learning foreign languages. *PloS One* 11, no. 6: e0158409. DOI: 10.1371/journal.pone.0158409.
- Bisson, M.J., W.J.B. Van Heuven, K. Conklin and R.J. Tunney. 2014. The role of repeated exposure to multimodal input in incidental acquisition of foreign language vocabulary. *Language Learning* 64, no. 4:

855–77. DOI: [10.1111/lang.12085](https://doi.org/10.1111/lang.12085).

Bravo, M.C.C. 2008. Putting the reader in the picture. Screen translation and foreign-language learning. PhD diss, University Rovira i Virgili and University of Algarve.

Bruton, A., M. García López and R. Esquiliche Mesa. 2011. Incidental L2 vocabulary learning: an impracticable term? *TESOL Quarterly* 45, no. 4: 759–68. DOI: [10.5054/tq.2011.268061](https://doi.org/10.5054/tq.2011.268061).

Chang, A.C.S. and J. Read. 2006. The effects of listening support on the listening performance of EFL learners. *TESOL Quarterly* 40, no. 2: 375–97.

Chen, Y., Y. Liu and A.G. Todd. 2018. Transient but effective? captioning and adolescent EFL learners' spoken vocabulary acquisition. *English Teaching & Learning* 42, no. 1: 25–56. DOI: [10.1007/s42321-018-0002-8](https://doi.org/10.1007/s42321-018-0002-8)

Chung, J.M. 1996. The effects of using advance organizers and captions to introduce video in the foreign language classroom. *TESL Canada Journal* 14, no. 1: 61–5.

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, no. 3: 295–308.

Danan, M. 2004. Captioning and subtitling: undervalued language learning strategies. *Meta: Journal Des Traducteurs* 49, no. 1: 67–77. DOI: [10.7202/009021ar](https://doi.org/10.7202/009021ar).

Dörnyei, Z. 2009. *The Psychology of Second Language Acquisition*. Oxford: Oxford University Press.

D'Ydewalle, G. and M. Van de Poel. 1999. Incidental foreign-language acquisition by children watching subtitled television programs. *Journal of Psycholinguistic Research*, 28, no. 3: 227–44. DOI: [10.1023/A:1023202130625](https://doi.org/10.1023/A:1023202130625)

Elley, W.B. 1989. Vocabulary acquisition from listening to stories. *Reading Research Quarterly* 24, no. 2: 174–87. European Commission. 2017. *Media use in the European Union*. European Union, Standard Eurobarometer 88 Project. DOI: [10.2775/116707](https://doi.org/10.2775/116707)

Frumuselu, A.D., S. De Maeyer, V. Donche and M.M. Gutiérrez Colon-Plana. 2015. Television series inside the EFL classroom: bridging the gap between teaching and learning informal language through subtitles.' *Linguistics and Education* 32: 107–17. DOI: [10.1016/j.linged.2015.10.001](https://doi.org/10.1016/j.linged.2015.10.001).

Gesa, F. and I. Miralpeix. 2018. Enhancing foreign language learning by means of multimodal input. The case of subtitled TV series and young learners. Paper presented at the 2018 Conference on Technological Innovation for Specialized Linguistic Domains, Ghent, May 24–26.

Guillory, H.G. 1998. The effects of keyword captions to authentic French video on learner comprehension. *Calico Journal* 15, no. 1–3: 89–108.

Horst, M., T. Cobb and P. Meara. 1998. Beyond a clockwork orange: acquiring second language vocabulary through reading. *Reading in a Foreign Language* 11, no. 2: 207–23.

- Hsu, C.K., J.G. Hwang, Y.T. Chang and C.K. Chang. 2013. Effects of video caption modes on English listening comprehension and vocabulary acquisition using handheld devices. *Journal of Educational Technology & Society*, 16, no. 1: 403–14. ISSN1436-4522
- Hulstijn, J.H. 2003. Incidental and intentional learning. In *Handbook of Second Language Acquisition* ed. C.J. Doughty and M.H. Long, 349–81. Malden, MA: Wiley-Blackwell.
- Hulstijn, J.H. 2013. Incidental learning in second language acquisition. In *The Encyclopedia of Applied Linguistics*, ed. C.A. Chapelle, vol. 5, 2632–40. Chichester: Wiley-Blackwell. DOI: [10.1002/9781405198431.wbeal0530](https://doi.org/10.1002/9781405198431.wbeal0530)
- Hummel, K.M. 2010. Translation and short-term L2 vocabulary retention. Hindrance or help? *Language Teaching Research* 14: 61–74. DOI: [10.1177/1362168809346497](https://doi.org/10.1177/1362168809346497).
- Jones, L. 2004. Testing L2 vocabulary recognition and recall using pictorial and written test items. *Language Learning & Technology* 8, no. 3: 122–43. ISSN 1094-3501
- Khan, N., J. Kasdar, M. Melvin, R. Blomquist, E. Huang and J. McEwen. 2015. *Fresh off the Boat* [TV Series]. Los Angeles, CA: ABC
- Koolstra, C.M. and J.W.J. Beentjes. 1999. Children's vocabulary acquisition in a foreign language through watching sub-titled television programs at home. *Educational Technology Research and Development* 47, no. 1: 51–60. doi:[10.1007/bf02299476](https://doi.org/10.1007/bf02299476).
- Kuppens, A.H. 2010. Incidental foreign language acquisition from media exposure. *Learning, Media and Technology* 35, no. 1: 65–85. doi:[10.1080/17439880903561876](https://doi.org/10.1080/17439880903561876).
- Laufer, B. 2005. Focus on form in second language vocabulary learning. *EUROSLA Yearbook* 5, no. 1: 223–50. Lekkai, I. 2014. Incidental foreign-language acquisition by children watching subtitled television programs. *TOJET: The Turkish Online Journal of Educational Technology* 13, no. 4: 81–7.
- Lindgren, E. and C. Muñoz. 2013. The influence of exposure, parents, and linguistic distance on young European learners' foreign language comprehension. *International Journal of Multilingualism* 10, no. 1: 105–29. DOI: [10.1080/14790718.2012.679275](https://doi.org/10.1080/14790718.2012.679275).
- Lwo, L. and M.C. Lin. 2012. The effects of captions in teenagers' multimedia L2 learning. *ReCALL* 24, no. 2: 188–208. DOI: [10.1017/s0958344012000067](https://doi.org/10.1017/s0958344012000067).
- Malone, J. 2018. Incidental vocabulary learning in SLA. *Studies in Second Language Acquisition*, 1–25. DOI: [10.1017/s0272263117000341](https://doi.org/10.1017/s0272263117000341).
- Matielo, R., T. Collet and R.C.S.F. D'Ely. 2013. The effects of interlingual and intralingual subtitles on vocabulary learning by Brazilian EFL learners: an exploratory study. *Intercâmbio. Revista do Programa de Estudos Pós-Graduados em Linguística Aplicada e Estudos da Linguagem* 27: 83–99. ISSN2237-759x.
- Matielo, R., R.C.S.F. DEly and L. Baretta. 2015. The effects of interlingual and intralingual subtitles on second language learning/acquisition: a state-of-the-art review. *Trabalhos Em Linguística Aplicada* 54, no. 1: 161–82.

DOI: [10.1590/ 0103-18134456147091](https://doi.org/10.1590/0103-18134456147091).

Meara, P. and J. Milton. 2003. *X-lex: The Swansea Levels Test*. Newbury: Express Publishing.

Mohd Jelani, N.A. and F. Boers. 2018. Examining incidental vocabulary acquisition from captioned video. *Approaches to Learning, Testing, and Researching L2 Vocabulary ITL – International Journal of Applied Linguistics* 169, no. 1: 169–90. DOI:[10.1075/itl.00011.jel](https://doi.org/10.1075/itl.00011.jel).

Montero-Perez, M., E. Peters, G. Clarebout and P. Desmet. 2014. Effects of captioning on video comprehension and incidental vocabulary learning. *Language Learning & Technology* 18, no. 1: 118–41. DOI: [10125/44357](https://doi.org/10.10125/44357)

Montero-Perez, M., W. Van Den Noortgate and P. Desmet. 2013. Captioned video for L2 listening and vocabulary learning: a meta-analysis. *System* 41, no. 3: 720–39. DOI: [10.1016/j.system.2013.07.013](https://doi.org/10.1016/j.system.2013.07.013).

Muñoz, C. 2017. The role of age and proficiency in subtitle reading. An eye-tracking study. *System* 67: 77–86. DOI: [10.1016/ j.system.2017.04.015](https://doi.org/10.1016/j.system.2017.04.015)

Muñoz, C., T. Cadierno and I. Casas. 2018. Different starting points for English language learning: a comparative study of Danish and Spanish young learners. *Language Learning* 68. DOI: [10.1111/lang.12309](https://doi.org/10.1111/lang.12309)

Naghizadeh, M. and T. Darabi. 2015. The impact of bimodal, Persian and no-subtitle movies on Iranian EFL learners' L2 vocabulary learning. *Journal of Applied Linguistics and Language Research* 2, no. 2: 66–79. ISSN2376-760x

Nation, P. 2006. How large a vocabulary is needed for reading and listening? *Canadian Modern Language Review*, 63, no. 1: 59–82. DOI: [10.3138/cmlr.63.1.59](https://doi.org/10.3138/cmlr.63.1.59)

Nation, P. 2007. The four strands. *International Journal of Innovation in Language Learning and Teaching*, 1, no. 1: 2–13. DOI: [10.2167/illt039.0](https://doi.org/10.2167/illt039.0)

Nation, P. 2012. The BNC/COCA Word Family Lists. Document bundled with Range program with BNC/COCA lists, 25.

Nation, P. 2015. Principles guiding vocabulary learning through extensive reading. *Reading in a Foreign Language* 27, no.1: 136–45. <http://nflrc.hawaii.edu/rfl>

Nation, P. and Heatley, A. 2002. Range: a program for the analysis of vocabulary in texts [software]. <https://www.victoria.ac.nz/lals/about/staff/paul-nation>

Neuman, S.B. and P. Koskinen. 1992. Captioned television as comprehensible input: effects of incidental word learning from context for language minority students. *Reading Research Quarterly* 27, no. 1: 94–106. DOI: [10.2307/747835](https://doi.org/10.2307/747835).

Peters, E. 2018. The effect of out-of-class exposure to English language media on learners' vocabulary knowledge. *Approaches to Learning, Testing, and Researching L2 Vocabulary ITL - International Journal of Applied Linguistics* 169, no. 1: 142–68. doi:[10.1075/itl.00010.pet](https://doi.org/10.1075/itl.00010.pet).

- Peters, E., E. Heynen and E. Puimège. 2016. Learning vocabulary through audiovisual input: the differential effect of L1 subtitles and captions. *System* 63: 134–48. DOI: [10.1016/j.system.2016.10.002](https://doi.org/10.1016/j.system.2016.10.002).
- Peters, E. and S. Webb. 2018. Incidental vocabulary acquisition through viewing L2 television and factors that affect learning. *Studies in Second Language Acquisition* 1–27. DOI: [10.1017/s0272263117000407](https://doi.org/10.1017/s0272263117000407).
- Pujadas, G. 2019. Language learning through extensive TV viewing. A study with adolescent EFL learners. PhD diss, Universitat de Barcelona.
- Pujadas, G. and C. Muñoz. 2017. Learning through subtitles. Learners' preferences and task perception. Paper presented at the 2017 International Conference on Task-Based Language Teaching, Barcelona, 19–21 April.
- Pujadas, G. and C. Muñoz. (in press) Examining adolescent EFL learners' TV viewing comprehension through captions and subtitles. *Studies in Second Language Acquisition*. First view at: <https://doi.org/10.1017/S0272263120000042>
- Rice, M.L., A.C. Huston, R. Truglio and J.C. Wright. 1990. Words from 'Sesame Street': learning vocabulary while viewing. *Developmental Psychology* 26, no. 3: 421–8. DOI: [10.1037//0012-1649.26.3.421](https://doi.org/10.1037//0012-1649.26.3.421).
- Rodgers, M.P.H. 2013. English language learning through viewing television: an investigation of comprehension, incidental vocabulary acquisition, lexical coverage, attitudes, and captions. PhD diss, Victoria University of Wellington.
- Rodgers, M.P.H. 2018. The images in television programs and the potential for learning unknown words. *Approaches to Learning, Testing, and Researching L2 Vocabulary ITL - International Journal of Applied Linguistics* 169, no.1: 191–211. DOI: [10.1075/itl.00012.rod](https://doi.org/10.1075/itl.00012.rod).
- Rodgers, M.P.H. and S. Webb. 2011. Narrow viewing: the vocabulary in related television programs. *TESOL Quarterly* 45, no. 4: 689–717. DOI: [10.5054/tq.2011.268062](https://doi.org/10.5054/tq.2011.268062)
- Schmitt, N. 2008. Review article: instructed second language vocabulary learning. *Language Teaching Research* 12, no. 3: 329–63. DOI: [10.1177/1362168808089921](https://doi.org/10.1177/1362168808089921)
- Schmitt, N. 2010. *Researching Vocabulary: A Vocabulary Research Manual*. Basingstoke: Palgrave Macmillan.
- Sockett, G. 2014. *The Online Informal Learning of English*. Basingstoke: Palgrave Macmillan.
- Stewart, M.A. and I. Pertusa. 2004. Gains to language learners from viewing target language closed-captioned films. *Foreign Language Annals* 37, no. 3: 438–42. DOI: [10.1111/j.1944-9720.2004.tb02701.x](https://doi.org/10.1111/j.1944-9720.2004.tb02701.x).
- Sydorenko, T. 2010. Modality of input and vocabulary acquisition. *Language Learning & Technology* 14, no. 2: 50–73. DOI: [10.125/44214](https://doi.org/10.125/44214)
- Vanderplank, R. 1990. Paying attention to the words: Practical and theoretical problems in watching television programmes with uni-lingual (CEEFAX) sub-titles. *System* 18, no. 2: 221–34.

Vanderplank, R. 2010. Déjà vu? A decade of research on language laboratories, television and video in language learning. *Language Teaching* 43, no. 1: 1–37. DOI: [10.1017/S0261444809990267](https://doi.org/10.1017/S0261444809990267)

Vanderplank, R. 2016. ‘Effects of’ and ‘effects with’ captions: how exactly does watching a TV programme with same- language subtitles make a difference to language learners? *Language Teaching* 49, no. 2: 235–50. DOI: [10.1017/s0261444813000207](https://doi.org/10.1017/s0261444813000207).

Van Zeeland, H. and N. Schmitt. 2013. Lexical coverage in L1 and L2 listening comprehension: the same or different from reading comprehension? *Applied Linguistics*, 34, no. 4: 457–79. DOI: [10.1093/applin/ams074](https://doi.org/10.1093/applin/ams074)

Webb, S. 2010. Pre-learning low-frequency vocabulary in second language television programmes. *Language Teaching Research* 14, no. 4: 501–15. DOI: [10.1177/1362168810375371](https://doi.org/10.1177/1362168810375371)

Webb, S. and A.C.S. Chang. 2015a. How does prior word knowledge affect vocabulary learning progress in an extensive reading program? *Studies in Second Language Acquisition*, 37, 651–75. DOI: [10.1017/S0272263114000606](https://doi.org/10.1017/S0272263114000606)

Webb, S. and P. Nation. 2017. *How Vocabulary is Learned*. Oxford: Oxford University Press.

Webb, S. and M.P.H. Rodgers. 2009a. Vocabulary demands of television programs. *Language Learning* 59, no. 2: 335–66. DOI: [10.1111/j.1467-9922.2009.00509.x](https://doi.org/10.1111/j.1467-9922.2009.00509.x).

Webb, S. and M.P.H. Rodgers. 2009b. The lexical coverage of movies. *Applied Linguistics* 30, no. 3: 407–27. DOI: [10.1093/applin/amp010](https://doi.org/10.1093/applin/amp010).

Winke, P., S. Gass and T. Sydorenko. 2010. The effects of captioning videos used for foreign language listening activities. *Language Learning & Technology* 14, no. 1: 65–86. DOI: [10.125/44203](https://doi.org/10.125/44203)

Zarei, A.A. 2008. The effect of bimodal, standard, and reversed subtitling on L2 vocabulary recognition and recall. *Pazhuhesh-e Zabanha-ye Khareji* 49: 65–84.

Zarei, A.A. and Z. Rashvand. 2011. The effect of interlingual and intralingual, verbatim and nonverbatim subtitles on L2 vocabulary comprehension and production. *Journal of Language Teaching and Research* 2, no. 3: 618–25. DOI: [10.4304/jltr.2.3.618-625](https://doi.org/10.4304/jltr.2.3.618-625).