

Extensive Viewing in Young English as a Foreign Language Learners: Do Aptitude and Vocabulary Size Influence Vocabulary Learning?

Audiovisual input has been shown to promote vocabulary acquisition in foreign language learners. However, there is limited research on the effects of extensive viewing in a classroom context with young beginner learners of English. This study investigates foreign language vocabulary learning through explicit teaching and additional viewing of subtitled TV series, as well as the role that second language (L2) vocabulary size and language aptitude may play. A teaching intervention lasting one academic year was implemented in primary school. Weekly, two groups of beginner learners of English enrolled in grade 6 and aged 11 ($N=47$) were taught a set of target words. One group was additionally exposed to the target vocabulary through an episode of a TV series with subtitles in Spanish, their native language. Gains in form and meaning recall of the target vocabulary were measured. Participants' L2 vocabulary size and language aptitude were also assessed at the beginning of the intervention. Multivariate tests showed that explicit teaching accounted for most vocabulary learning, although viewing the TV series significantly benefited learners at two testing points. L2 vocabulary size contributed to explaining lexical gains, while language aptitude did not have such a significant effect.

Keywords: extensive viewing; language aptitude; L2 vocabulary learning; L2 vocabulary size; vocabulary teaching; young learners

1. Introduction

To date, research has shown that original version (OV) films and television (TV) programs are popular resources for language learners (De Wilde et al., 2020, 2022; Muñoz, 2020). In this context, it is important to connect extramural exposure with classroom instruction. Likewise, integrating informal learning practices, such as OV TV viewing, into more formal instructional experiences could enhance incidental vocabulary learning opportunities (Teng, 2025). Moreover, recent literature also indicates a positive relationship between OV TV and learners' vocabulary attainment in the foreign language (FL) (e.g., Puimège & Peters, 2019). Nevertheless, as noted by Vanderplank (2016) and Montero Perez (2022), there is a lack of classroom-based studies on extensive viewing, which refers to the regular use of second language (L2) TV for learning purposes (Rodgers, 2016; Webb, 2015). Additionally, most existing research has focused on adolescents or adults (Gesa & Miralpeix, 2023; Rodgers & Webb, 2020; Vu et al., 2023), highlighting the need for more studies involving young learners. For instance, Reynolds et al. (2022), in a recent scoping review on vocabulary acquisition through video viewing, found that only six of the 34 studies (17.65%) that met the inclusion criteria focused on primary school learners. This underscores the need for more research involving young learners. This idea also aligns with the necessity to further investigate how children acquire an FL in classroom settings, as the current body of knowledge largely stems from studies with older, cognitively more mature populations, despite the increasing number of schools introducing FLs at an early age (Luquin & García Mayo, 2021).

It is important to note that L2 vocabulary acquisition by children in input-limited contexts differs from that by adults. Studies usually show that adult learners tend to progress faster than younger learners in instructed settings in terms of vocabulary

breadth, and that early starters do not seem to overtake late starters (e.g., Cadierno et al., 2020; Cenoz, 2002; Miralpeix, 2008). This suggests that vocabulary learning from extensive viewing may follow these similar patterns observed in other areas of vocabulary research, and that results found with teenagers or adults may not be directly applicable to younger learners.

2. Literature Review

2.1 Incidental Vocabulary Learning from Viewing in Young Learners

Although several studies have investigated extensive viewing –see Montero Perez et al.’s (2013) meta-analysis and Montero Perez’s (2022) state-of-the-art article–, there is scarce research focusing on young learners (Reynolds et al., 2022). Some studies with this population (e.g., Kuppens, 2010; Rice et al., 1990) have examined whether at-home TV viewing facilitates vocabulary development, while others (e.g., Alexiou, 2015; Teng, 2019a) have assessed the effects of first language (L1) subtitles and captions (i.e., L2 subtitles), or compared audiovisual input to other input modalities. However, to our knowledge, no studies have explored the potential benefits of extensive viewing as a classroom activity when integrated with traditional teaching practices.

In studies analyzing at-home TV viewing, Rice et al. (1990) tracked children aged three and five over two years (until the ages of five and seven, respectively), and collected TV viewing diaries which were compared to results from picture-matching meaning recognition vocabulary tests. Their findings indicated a correlation between children’s viewing habits and their scores on meaning recognition tests, particularly for three-year-olds. Similarly, Kuppens (2010) surveyed 374 grade six (G6) students about their frequency of target language (TL) media use at home, including (un)subtitled TV programs, music, computer games, etc. They also completed two oral translation tests

(L1-TL, TL-L1). The study found that watching subtitled audiovisual material was directly linked to higher scores on both translation tests, more so than any other type of input tapped in the survey. Wright et al. (2001) also found that the frequency of exposure to children's programs positively impacted a range of tests measuring school readiness and general academic achievement. However, this positive relationship has not always been observed. For instance, Patterson (2002) did not find a correlation between the frequency of TV viewing at home and bilingual toddlers' vocabulary size (VS) in English and Spanish, attributing the lack of significant interaction to the passive nature of TV viewing and suggesting that vocabulary learning may require more active engagement.

Experimental studies conducted in language classes have demonstrated that OV TV viewing can significantly enhance vocabulary learning among (very) young learners. For instance, Alexiou (2015) found that Greek preschoolers who repeatedly watched four five-minute episodes of *Peppa Pig* (80 minutes in total) were able to recognize one third of the aural word forms on which they were tested, despite the absence of a pre-test. Additionally, OV TV viewing has been reported to be more beneficial than other modalities of input. Neuman and Koskinen (1992) compared the effects of (un)captioned TV, reading-while-listening, and textbook use on incidental vocabulary acquisition among seventh and eighth graders. Their results showed that those exposed to captioned TV significantly outperformed the other groups, which the authors attributed to the simultaneous presentation of information through multiple modes.

The question of whether videos should be presented with or without textual support has been examined using different materials and tests. Most research suggests

that subtitles benefit children's vocabulary development (Arifani, 2020; Hariffin & Said, 2019; Koolstra & Beentjes, 1999), although some studies have reported opposite results (Linebarger et al., 2013). Of particular interest is the longitudinal three-month study by Chen et al. (2018), in which eighth graders were exposed to ten episodes of an animated TV series. Half of the participants viewed the episodes with captions, while the other half watched them without textual support. Learners were subsequently divided into three subgroups according to their proficiency level (high-, mid- and low-achievers). The results indicated a positive effect of captions on two recognition tests administered after each of the ten episodes, targeting 70 words in total (seven per episode). Vocabulary gains were also positively mediated by participants' TL proficiency, with high-achievers benefiting more from the experience than low-achievers.

Regarding the language and the type of the on-screen text, Pujadas and Muñoz (2019) tracked the vocabulary development of grade eight English as a foreign language (EFL) learners throughout an academic year, but did not find significant differences between captions and L1 subtitles. However, Casulleras (2023) found that fifth- and sixth-grade EFL learners benefited more from captions than from L1 subtitles when tested on word form recognition immediately after watching each of the 20 episodes they were exposed to over a five-month period. These findings contrast with previous one-off studies assessing meaning recognition, which indicated that children exhibited higher indices of vocabulary learning when exposed to L1 subtitles (e.g., d'Ydewalle & Pavakanun, 1997; d'Ydewalle & Van de Poel, 1999). Regarding captioning type, Teng and Cui (2024) investigated whether 12-year-old children learning English in Hong Kong benefitted more from full, keyword or no captions, compared to a control group

which only took the tests. The researchers selected two short storytelling videos, each lasting 12 minutes, and asked participants to provide English or Chinese definitions for the 35 target words (TW) and collocations on which they were tested. Statistical analyses for individual words revealed that the full caption condition resulted in the best performance, although all video-viewing conditions outperformed the control group. The authors attributed these results to the more detailed information provided by full captions, which participants may have used to process TWs more easily, compared to the more concise information provided by keyword captions, which may have been insufficient to facilitate vocabulary intake and uptake.

2.2 Vocabulary Pre-Teaching when Learning from Viewing in Young Learners

Although existing research has suggested a positive effect of TV viewing on children's vocabulary development, the gains derived from exposure to audiovisual input tend to be limited (e.g., Gesa & Miralpeix, 2022). To maximize learning, one strategy is to supplement viewing with intentional learning of the target vocabulary through teacher-led instruction, thereby combining both incidental and intentional approaches (Montero Perez et al., 2018; Schmitt, 2008; Webb, 2020). Vocabulary teaching is said to facilitate a relatively high number of encounters with TWs (Schmitt & Schmitt, 2020), and teachers' explanations of the target vocabulary have proved beneficial for vocabulary development (Lee & Lee, 2022). Furthermore, research has shown that, under similar conditions and comparable amounts of input, more words can be learned intentionally than incidentally (Barcroft, 2015; Laufer & Rozovski-Roitblat, 2015; Lindstromberg, 2020). In relation to video viewing, Webb (2010) suggested that pre-teaching words prior to extensive viewing enhances comprehension and vocabulary

learning, although Reynolds et al. (2022) found no significant differences between incidental and intentional learning of vocabulary through video viewing.

In this respect, advance organizers –introductory relevant information presented prior to a lecture (Ausubel, 1968)– have been shown to be beneficial in learning from video viewing (Chung, 1996, 2002; Yang, 2014). Various types of advance organizers have been used in research, including previewing comprehension questions (Jafari & Hashim, 2012), annotating (Babaei & Izadpanah, 2019), brainstorming (Li et al., 2019), and graphic organizers such as diagrams (Jiang & Grabe, 2007). One extensively investigated type is the presentation of key vocabulary, which has proven particularly advantageous (Chung, 1996, 2002; Elkhafaifi, 2005; Jafari & Hashim, 2012). Advance organizers have also been used with children learning a new language (e.g., Hanley et al., 1995; Rinehart & Welker, 1992; Wilberschied & Berman, 2004), yielding promising results.

More specifically, Severson (2017) and Pujadas and Muñoz (2019) concluded that teacher-led instruction serves as a positive reinforcement for L2 vocabulary learning from video viewing among young learners, also when repeated exposure to OV TV viewing was guaranteed. Similarly, Teng (2019a) examined incidental collocation learning with and without advance organizers (i.e., matching images to brief synopses of the videos watched) and different captioning techniques (full, keyword, and no captions) with 12-year-old children, who were exposed to four 10-minute YouTube videos. The study found that presenting advance organizers prior to viewing facilitated the acquisition of various aspects of collocation knowledge, including both receptive and productive knowledge of form and meaning, while also reducing cognitive load. In a related study, Teng (2022b) investigated whether advance organizers –using the same

strategy as in Teng (2019a)– facilitated vocabulary learning among 240 Chinese sixth graders learning English as an L2. Participants watched four short YouTube videos with different captioning techniques (glossed, full, and keyword captions) and were tested on 20 unknown words using a battery of post-tests assessing word form and meaning recognition, and word use. The results revealed that advance organizers enhanced all aspects of word knowledge, with their combination with glossed captioning proving to be the most effective.

Similarly, Avello (2023) exposed Chilean primary school students in grades four and five to 11 captioned episodes of *Charlie and Lola*, followed by activities focused on either meaning (e.g., multiple-choice questions assessing episode comprehension) or vocabulary construction (e.g., tasks requiring interaction with the target vocabulary). Before and after watching the ten-minute episodes, participants took a dictation test to assess word form recall and a multiple-choice meaning recognition test. The latter was also administered as a delayed test three weeks after the study ended. The analyses revealed that both types of post-viewing activities were equally effective, although construction-focused activities played a crucial role in word form recall. Hence, Avello (2023) concluded that "the use of audiovisual materials should be complemented by intentional activities that may boost learners' gains without interfering with the viewing process" (p. 353), particularly for recall knowledge.

Finally, Suter et al. (2024) examined the effects of pre- and post-viewing activities on learning FL vocabulary among 97 beginner young German speakers learning L2 French. Participants were exposed to three short episodes of a Netflix TV series (13-15 minutes on average) with on-screen text in French. For two of the episodes, they completed either a pre- or post-viewing activity that focused on the target

vocabulary, while the third episode was watched without any supplementary activity. After viewing each episode, learners completed a meaning recognition test, which closely mirrored the supplementary activities and assessed only the 17 TWs encountered in that episode. A one-week delayed test, covering all TWs ($N=51$), was also included in the design. The results revealed that incorporating pre- or post-viewing activities was essential for promoting vocabulary learning and short-term retention, with participants learning seven or eight more words when a pre- or post-task was completed. However, the difference between pre- and post-activities was negligible.

Pre-teaching of lexical items is not the only factor that may influence vocabulary learning from viewing. Recent research has shown that learner-related factors may also play a significant role (Montero Perez, 2022). Notably, prior vocabulary knowledge and language aptitude appear to be linked to vocabulary acquisition from video viewing (e.g., Suárez & Gesa, 2019).

2.3 Learner Factors Influencing Learning from Viewing: L2 Vocabulary Size and Language Aptitude

Research involving adult populations has demonstrated that prior vocabulary knowledge –often measured as VS– is a critical factor when analyzing vocabulary learning from OV input. A larger VS has been associated with greater gains in form and meaning recognition, as well as improved results on recall tests (Ahrabi Fakhr et al., 2021; Montero Perez, 2020; Montero Perez et al., 2014, 2015; Muñoz et al., 2023; Peters et al., 2016; Vu et al., 2023). For instance, Peters and Webb (2018) found that for every ten words known in the VS test, the likelihood of correctly answering vocabulary recognition and recall tests after watching a full-length documentary increased by 32%. Similarly, Peters et al. (2016) showed that each known word in the VS test improved the

chances of a correct response in vocabulary tests by 2% to 5%, while Montero Perez (2020) reported a 3% increase. Suárez and Gesa (2019) also found that VS significantly impacted word form and meaning learning after three months of sustained exposure to captioned TV series and explicit teaching. However, other studies on video viewing have reported less consistent results in meaning recall tests (Teng, 2022a), potentially due to the complex nature of these tests and their mismatch with the more passive activity of watching OV TV.

Despite this growing body of evidence, to the best of our knowledge, only three studies have analyzed the role of VS when learning from viewing among young learners. Casulleras (2023), who administered the first three bands of the Vocabulary Size Test (Nation & Beglar, 2007), found a moderate correlation between VS and form recognition scores when learners were exposed to L1 subtitles, but not to captions. This result was attributed to captions already facilitating the acquisition of word forms, reducing the need for reliance on VS. Additionally, Avello (2023) found that VS, measured with a picture-based meaning recognition test, was a strong predictor of written word form recall and form and meaning recognition. However, Teng and Cui (2024) found no significant impact of vocabulary breadth, measured using the updated version of the Vocabulary Levels Test (Webb et al., 2017), on the learning of individual words, with only 6% of the variance being explained by prior vocabulary knowledge. These contrastive findings suggest that further research is needed to understand how children's L2 VS, which is typically more limited compared to that of older populations, may (not) influence the learning of new words from audiovisual input.

Research into the role of aptitude in vocabulary learning, particularly through audiovisual input among young learners, remains scarce. Vocabulary learning has been

linked to working memory and phonological memory skills, which are oftentimes considered components of aptitude (Gathercole & Baddeley, 1993). In this respect, Teng (2023) found that phonological working memory, but not complex working memory, significantly predicted incidental vocabulary acquisition from captioned video viewing among 12-year-old learners exposed to four YouTube videos with different captioning techniques (full, glossed, and keyword captions). In addition, Teng and Cui (2024) found that both phonological and complex working memory facilitated individual word learning among primary school children learning English through captioned video viewing. In terms of language aptitude, Casulleras (2023) demonstrated its influence on word form recognition when TV series were presented with L1 subtitles, but not with captions. Among older and more proficient learners at the university level, aptitude was found to be relevant for word meaning learning in Suárez and Gesa (2019), although it was intertwined with participants' VS. Positive effects of aptitude on both word form and meaning recognition, as well as recall, have also been observed by Teng (2022a). Similarly, Muñoz et al. (2022) found that sound recognition ability, an aptitude component, was associated with meaning recognition scores when learning from captioned videos among upper-intermediate college learners.

3. The Present Study

Addressing existing gaps in the literature –such as the limited research on extensive viewing among young learners, the impact of vocabulary pre-teaching, and the role of learner factors–, the present study implements a teaching intervention over one academic year, divided into three academic terms, within a primary school classroom. The study also examines the potential influence of VS and language aptitude. During the intervention, two groups of students were taught a set of TWs

through formal instruction. However, one group was additionally exposed to 24 episodes of subtitled TV series (eight episodes per term), in which the target vocabulary appeared. Hence, this study aims to address the following research questions (RQ):

In beginner EFL young learners,

1. To what extent does sustained exposure to L1-subtitled TV series, combined with formal instruction, enhance vocabulary learning compared to formal instruction alone?
2. To what extent do VS and language aptitude influence vocabulary learning from additional extensive viewing?

3.1 Participants

A total of 47 participants from two intact classes were selected for the study. One class ($n=25$) was assigned to the experimental group (EG), which received formal instruction on the target vocabulary and was also exposed to the TWs through the TV series. The other class ($n=22$) was assigned to the comparison group (CG), which received only formal instruction on the target vocabulary. All participants were Catalan / Spanish bilinguals studying EFL in a semi-private school, and they had a beginner level of the TL (A1-A2 following the Common European Framework of Reference for languages). They were all in G6, the final year of primary education, and were 11 years old (mean age: 10.98 years, $SD=.33$). The mean VS was 1,361 (EG) and 1,418 words (CG), as measured by the X_Lex test (Meara, 2005). Additionally, their average reading speed in Spanish –the same language used for the subtitles– was 155 words per minute, which allowed them to read the subtitles effectively, as the typical presentation speed for subtitles ranges from 130 to 160 words per minute (Díaz Cintas, 2012).

3.2 Instruments

Different instruments were used in the intervention: some were specifically designed for the study, including vocabulary pre- and post-tests, and vocabulary pre- and post-tasks. Others comprised authentic materials, such as TV series, and validated tests, such as the X_Lex test and the Modern Language Aptitude Test-Elementary Catalan.¹

3.2.1 TV Series and Target Vocabulary. The first 16 episodes of season one of *The Suite Life of Zack and Cody* (Kallis et al., 2005) were used in the first and second terms, and eight episodes of season one of *Wizards of Waverly Place* (Greenwald et al., 2007) in the third one. These two TV series were selected since they were aimed at pre-teens, were unseen by participants, were not being aired in Spain or in any streaming platform at the time of data collection, and the lexical profiles of the episodes selected showed that the TV series were adequate in terms of coverage at this proficiency level (see Table 1). Participants in the EG were exposed to 24 episodes in total (eight each term), totaling eight hours thirty-six minutes. Episodes were shown in English (L2) audio and Spanish (L1) subtitles, given the aid that L1 subtitles are thought to offer to beginner learners (Caimi, 2006).

As seen in Table 1, a corpus analysis of the episodes showed that 95% lexical coverage was reached at the 2K-3K levels, indicating that all the episodes selected were very similar as regards lexical difficulty and that L1 subtitles were necessary given the EG learners' VS (1,361 words) (Webb & Rodgers, 2009). A set of comprehension tests² administered after the viewing of each episode showed that comprehension was adequate, reaching levels above 65% in many episodes. Moreover, none of the participants reported having problems with understanding the storyline of the TV series in a subsequent questionnaire administered at the end of the intervention.

Table 1*Lexical Characteristics and Average Length of the Episodes Selected*

Term	1K coverage (in %)	2K coverage (in %)	3K coverage (in %)	Tokens	Types	Average length
Term 1	91.11	94.85	95.89	20,106	2,484	21' 27''
Term 2	91.18	94.82	95.83	21,848	2,722	21' 27''
Term 3	91.71	95.58	96.28	23,873	2,436	21' 36''

Five TWs were selected from each episode, so students were pre- and post-tested on a total of 120 TWs (40 per term). Various word variables were controlled, such as frequency of occurrence in the episode and in the language, concreteness, and cognateness. More frequent words in the episode were prioritized since research on audiovisual input has shown that they are better learned (e.g., Peters & Webb, 2018). However, very high-frequency words in corpora were avoided whenever possible, as they were likely known by participants (Ellis, 2002). Additionally, there were more concrete rather than abstract TWs, as concreteness facilitates language learning (Walker & Hulme, 1999). Finally, cognates between English and Catalan / Spanish were not included (please see Appendix A for the list of TWs).

3.2.2 Vocabulary Pre- and Post-Tests. All vocabulary pre- and post-tests followed the same format, allowing for the calculation of lexical gains (Nation & Webb, 2011), and included all the TWs for a given term ($n=40$). Four practice items were included at the beginning, with the rest of the TWs presented randomly. Learners listened to an audio file recorded by an English L1 speaker in which each TW form was read aloud twice. Participants had to write down the forms in English and provide the Catalan / Spanish translation or definition.

The tests were designed to measure written form recall (involving word form recognition and later transcription) and meaning recall. Thus, they analyzed distinct types of knowledge: receptive knowledge of the spoken form (not directly assessed by the tests but needed to complete the task successfully), productive knowledge of the written form, and passive recall (Laufer & Goldstein, 2004; Webb, 2007). The tests proved to be reliable, with high discriminatory power: Cronbach's alpha for relative gains for form was $\alpha=.875$, while for meaning it was $\alpha=.895$ (Tavakol & Dennick, 2011).

3.2.3 Vocabulary Pre- and Post-Tasks. All participants completed the vocabulary pre- and post-tasks in each session. The pre-task aimed to pre-teach the TWs to students and consisted of one focus-on-forms exercise in which TWs were the object of study (Laufer, 2006). Different types of tasks were created (e.g., crosswords, matching words with pictures / definitions, word searches, etc.) and extra help was provided (e.g., the first letter of the TW was circled in the word searches) to ensure they were not too difficult for G6 students. In the vocabulary post-task, participants listened to an audio file in which each TW form was read aloud twice, wrote down the English form of the words, and then selected the best Spanish translation out of five possibilities plus an 'I don't know' option –format adapted from Rodgers and Webb (2020). Thus, the post-task tapped into written form recall (prompted by the aural form of the words) and meaning recognition.

3.2.4 Vocabulary Size Test: X_Lex. Participants took the X_Lex v2.05 test (Meara, 2005) to check their receptive VS. X_Lex measures knowledge of word families within the 1-5K frequency band range. Testees are presented with a set of words (one at a time) and they respond whether they know their meaning or not. A

number of pseudo-words is also included to adjust scores for guessing. Participants who claimed to know six or more pseudo-words were excluded from the sample, as their results would not have been reliable (Zhang et al., 2019).

3.2.5 Aptitude Test: MLAT-EC. The MLAT-EC (Suárez, 2010) is the Catalan version of the MLAT-Elementary (Carroll & Sapon, 1967). Designed for learners in grades three to seven, the test consists of four main sections: ‘Hidden Words’ (30 items) evaluates phonetic coding ability, ‘Matching Words’ (30 items) assesses grammatical sensitivity, ‘Rhyming Words’ (38 items) focuses on rote learning ability, and ‘Number Learning’ (25 items) measures inductive language learning ability and rote memory. The MLAT-EC has been successfully validated in research with similar populations (Rispa, 2021), and its internal consistency in the present study was found to be acceptable (Cronbach’s $\alpha=.768$).

3.3 Procedure

The teaching intervention spanned one academic year and was divided into three terms. It was integrated into the school’s English curriculum, but it accounted for only one of the three one-hour English classes participants had each week, with the remaining two classes dedicated to following the standard curriculum and learning other aspects of the English language, including new vocabulary unrelated to the TWs. At the start of the academic year, the school board approved the experiment, and written informed consent was obtained from participants’ parents. Next, all learners completed the VS test (individually and computer-based) and the MLAT-EC (as a paper-and-pencil class activity).

Then, the same procedure was repeated every term. At the beginning, all participants took the vocabulary pre-test, which covered all the TWs they would

encounter during that term. One week after the pre-test, viewing sessions began and were held once a week for eight weeks, until the end of the term. At the start of each 55-minute session, both EG and CG completed the vocabulary pre-task, which was corrected immediately afterwards. The pre-tasks were collected by the teacher to minimize the likelihood of participants using them to study the TWs at home.

Next, only those allocated to the EG watched an episode from the TV series featuring the TWs. The episode was projected onto the classroom whiteboard with Spanish subtitles. Meanwhile, the CG, which did not view the TV series, followed the regular school curriculum, engaging in activities unrelated to the TWs and not involving video viewing. These activities did not provide the CG with additional exposure to the TWs, and no further explicit vocabulary instruction was given.

At the end of each session, all participants completed the vocabulary post-task. Answer sheets were immediately collected and corrected by the researcher to prevent participants from using them in future tests. At the end of each term, all participants took the vocabulary post-test, which assessed their knowledge of the 40 TWs covered during that term.

3.4 Scoring and Analysis

3.4.1 Pre- and Post-Test Scoring and Lexical Gains. For a TW form to be considered correct, strict spelling criteria were applied, following Webb (2007) and Pujadas and Muñoz (2023). Regarding TW meanings, acceptable responses included L1 translations, L1/L2 synonyms, or definitions. For polysemous words (only eight words according to the Oxford Learner's Dictionary), participants were asked to provide all the meanings they were familiar with. However, only the meaning presented in the TV series and covered in the vocabulary tasks was accepted.

Both absolute and relative gains were computed (Nation & Webb, 2011).

However, only relative gains were used in the statistical analysis. Relative gains account for learners' knowledge before the intervention, provide a more accurate measure of vocabulary learning (Horst et al., 1998), and have been widely used in recent video viewing research (e.g., Rodgers & Webb, 2020). The formula for relative gains is:

$$\text{Relative gains} = \frac{N \text{ of TWs learned}}{N \text{ of items tested} - N \text{ of TWs known}} \times 100$$

where, 'TWs learned' = N of items answered incorrectly on the pre-test but correctly on the post-test, and 'TWs known' = N of items answered correctly on both the pre- and post-test. Finally, the average gains for word form and meaning across the three terms were calculated.

3.4.2 Vocabulary Size and Language Aptitude Tests. For the VS test, the adjusted score, which accounts for guessing, was used in the analysis, with a maximum possible score of 5,000 word families. For the aptitude test, each question was worth one point, making the maximum score 123 points. The total aptitude score was used in the analysis because it is a good indicator of learners' overall ability to learn new FLs (Skehan, 1991).

3.4.3 Statistical Analysis. First, preliminary analyses were conducted to ensure that there were no significant differences between the EG and the CG in terms of their VS and aptitude scores. To do so, Mann-Whitney U tests were used for this purpose, as the data for the MLAT-EC were not normally distributed and the number of participants was limited.

To address the first RQ, two linear mixed models (LMM) were constructed, one for word form and one for word meaning. The meaning scores were log-transformed prior to model fitting to normalize their distribution. Both models had the same

structure, incorporating Time (Terms 1, 2, and 3) to analyze lexical gains longitudinally and Group (EG vs. CG) to examine differences between conditions. Additionally, the Time \times Group interaction was included to explore how effects varied across groups throughout the intervention. The models also included by-participant and by-item random intercepts where applicable. Compound symmetry was selected as the covariance type to ensure equal variances and covariances across all participants.

For the second RQ, data from the EG were analyzed, as they were the only participants exposed to the viewing sessions. Two-tailed Pearson product-moment correlations and Spearman's rank correlation coefficients were computed between the average of word forms and meanings learned, VS, and MLAT-EC total scores. Additionally, two hierarchical multiple regressions (one for word form and one for word meaning) were performed with relative gains as the dependent variables. VS was entered at stage one and language aptitude at stage two, based on previous research indicating that VS has a stronger relationship with vocabulary learning from video viewing in adult populations (Muñoz et al., 2022).

4. Results

A Mann-Whitney U test ($U=284.5$, $z=46.87$, $p=.839$) revealed no significant differences between the EG and the CG in terms of VS scores, nor in their language aptitude, as shown by another Mann-Whitney U test ($U=238.5$, $z=39.53$, $p=.613$).

4.1 RQ1 - Vocabulary Learning from Video Viewing and Teaching

The descriptive statistics for the relative gains computed for the two groups are presented in Table 2.³ As shown, the EG consistently achieved higher scores than the CG, although there was considerable individual variability, as evidenced by the high

standard deviations. This variability suggests that while some participants in the EG learned most TWs, others experienced minimal gains.

The LMM for word form learning revealed a significant main effect for Time, $F(2, 114)=4.325, p=.015, \eta_p^2=.071$, indicating a medium effect size (Cohen, 1988; Richardson, 2011). Pairwise contrasts showed no significant differences in the evolution of relative gains within the EG. However, the CG demonstrated significantly greater gains in Term 1 compared to Term 2, $t(114)=2.406, p=.018, 95\% \text{ CI } [.89, 9.21], d=.59$, and compared to Term 3, $t(114)=2.291, p=.024, 95\% \text{ CI } [.65, 8.92], d=.64$, both indicating small size effects. Regarding the Group effect, it was statistically significant as well, $F(1, 114)=5.098, p=.026, \eta_p^2=.043$, though with a small effect size. Pairwise contrasts showed significant differences between groups in Term 2, $t(114)=1.991, p=.049, 95\% \text{ CI } [.03, 12.18], d=.72$, and in Term 3, $t(114)=2.509, p=.014, 95\% \text{ CI } [1.61, 13.71], d=.78$, both with a medium effect size. The interaction between Time and Group was not statistically significant, $F(2, 114)=.942, p=.393, \eta_p^2=.016$, suggesting that the combined effect of Time and Group did not exceed the effects of the individual factors (see Figure 1 for a graphical representation of the results).

Table 2

Descriptive Statistics for Relative Gains for Form and Meaning

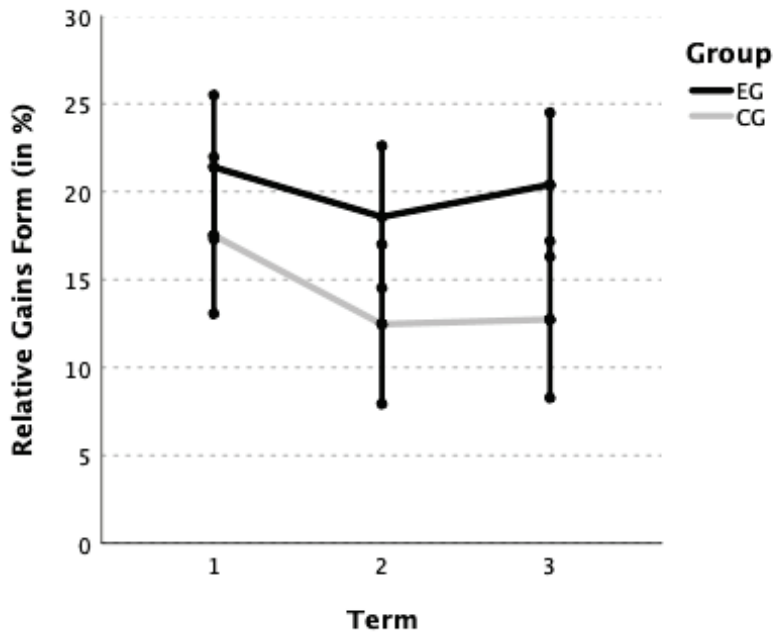
Group	Form*			Meaning*		
	<i>M</i>	<i>SD</i>	95% CI	<i>M</i>	<i>SD</i>	95% CI
Term 1						
EG (<i>n</i> =25)	20.82	10.75	[16.06, 25.59]	10.33	8.35	[6.63, 14.03]
CG (<i>n</i> =22)	18.39	10.42	[13.21, 23.57]	10.80	10.62	[5.52, 16.08]
Term 2						
EG (<i>n</i> =25)	18.08	10.49	[13.54, 22.61]	11.30	6.22	[8.61, 13.99]
CG (<i>n</i> =22)	13.07	7.30	[9.32, 16.82]	9.48	9.35	[4.68, 14.29]

	Term 3					
EG (<i>n</i> =25)	20.19	11.43	[15.12, 25.26]	11.70	7.05	[8.58, 14.83]
CG (<i>n</i> =22)	12.28	8.59	[8.01, 16.55]	8.47	5.95	[5.51, 11.43]

*All scores are shown in percentages.

Figure 1

Relative Gains for Word Form Learning Throughout the Intervention, According to Group

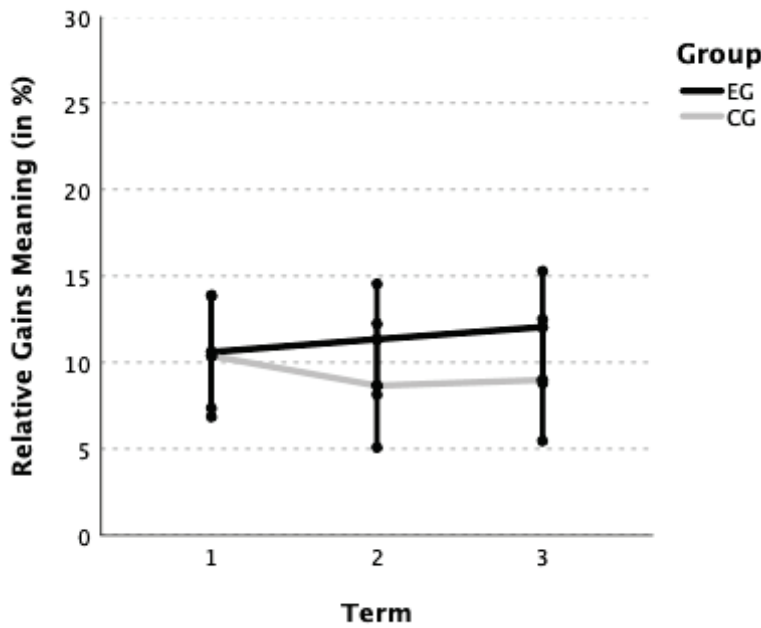


Regarding word meaning, the LMM indicated a non-significant main effect for Time, $F(2, 114)=1.774, p=.174, \eta_p^2=.030$. However, pairwise contrasts revealed that the relative gains experienced by the EG in Term 3 were significantly higher than those in Term 1, $t(114)=2.030, p=.045, 95\% \text{ CI } [.01, .70], d=.18$, although the effect size was very small. Group did not have a statistically significant impact on the learning of word meanings, $F(1, 114)=2.905, p=.091, \eta_p^2=.025$. Nonetheless, significant differences in favor of the EG in Term 2 were observed, $t(114)=2.173, p=.032, 95\% \text{ CI } [.05, 1.05], d=.23$, although the effect size was very small. Finally, the interaction between Time

and Group was not statistically significant, $F(2, 114)=1.257, p=.288, \eta_p^2=.022$ (see Figure 2 for a graphical representation of the results).

Figure 2

Relative Gains for Word Meaning Learning Throughout the Intervention, According to Group



4.2 RQ2 - Vocabulary Size and Language Aptitude in Vocabulary Learning from Video Viewing and Teaching

Regarding RQ2, it is important to note that the analysis was conducted only for participants in the EG. Table 3 presents the average gains for word form and word meaning over the academic year, as well as the VS and MLAT-EC scores for those participants with a complete dataset ($n=18$).

Table 3

Descriptive Statistics for Average Relative Gains for Form and Meaning, X_{Lex} , and MLAT-EC Scores

	<i>M</i>	<i>SD</i>	95% CI
Average word form gains*	18.37	9.99	[13.41, 23.35]

Average word meaning gains*	10.51	6.62	[7.22, 13.81]
X_Lex**	1,214	661	[885, 1,543]
MLAT-EC***	86.11	22.61	[74.87, 97.36]

*Relative gains scores are shown in percentages.

**The maximum score for X_Lex was 5,000 word families.

***The maximum score for MLAT-EC was 123 points.

Pearson and Spearman two-tailed correlations showed that VS was positively and moderately correlated with word form learning, $r=.567$, $p=.014$, 95% CI [.12, .81], as well as with word meaning, $\rho=.552$, $p=.017$, 95% CI [.10, .82], with both correlations demonstrating medium effect sizes (Plonsky & Oswald, 2014). Similarly, the total aptitude score was significantly correlated with word form learning, $\rho=.579$, $p=.012$, 95% CI [.14, .83], but it did not reach statistical significance with word meaning, $\rho=.419$, $p=.083$, 95% CI [-.07, .75], showing medium effect sizes in both cases.

The hierarchical multiple regression for word form learning showed that both VS and language aptitude accounted for 42.3% of the variance, with the model being statistically significant, $F(2, 15)=5.508$, $p=.016$. VS alone explained 32.1% of the variance, $F(1, 16)=7.594$, $p=.014$. However, adding language aptitude, which accounted for an additional 10.2%, did not lead to a significant change, $F(1, 15)=2.642$, $p=.125$. For word meaning, VS explained 24.7% of the variance in relative gains, $F(1, 16)=5.249$, $p=.036$. Adding language aptitude, which explained an additional 5.2%, did not result in a significant change either, $F(1, 15)=1.122$, $p=.306$. When aptitude was included, the model was not statistically significant, $F(2, 15)=3.205$, $p=.069$, although it accounted for a total of 29.9% of the variance (see Table 4 for the regression statistics).

Table 4*Summary of Hierarchical Regressions for Variables Predicting Relative Gains*

Variable	β	t	p	R	R^2	ΔR^2
Word form						
Step 1				.567	.322	.322
Vocabulary size	.567	2.756	.014			
Step 2				.651	.423	.102
Vocabulary size	.449	2.148	.048			
Language aptitude	.340	1.625	.125			
Word meaning						
Step 1				.497	.247	.247
Vocabulary size	.497	2.291	.036			
Step 2				.547	.299	.052
Vocabulary size	.412	1.788	.094			
Language aptitude	.244	1.059	.306			

5. Discussion

The aim of the present study was twofold: first, to determine whether additional exposure to TV series, in conjunction with formal instruction, was more effective for vocabulary learning compared to teacher-led instruction alone; second, to assess the role of VS and language aptitude in learning new vocabulary from video viewing and explicit teaching. The results indicated that additional exposure to TV series positively impacted word form learning, particularly in Terms 2 and 3, but did not significantly affect word meaning, except in Term 2. That said, most TWs were learned intentionally through explicit teaching. Furthermore, when learners were exposed to L1-subtitled TV series and teaching, prior L2 vocabulary knowledge appeared to have a stronger influence on learning outcomes than language aptitude.

5.1 L2 Vocabulary Learning from Video Viewing and Teaching

As anticipated, both the CG and the EG learned new words due to the deliberate attention to vocabulary items through focus-on-forms tasks. When learners are made

aware of target items, it often leads to learning (Schmidt, 2012). It is hence important to recognize that most of the learning occurred through these intentional activities (Lee & Lee, 2022; Schmitt, 2008), and differences between groups were not consistently evident throughout the intervention. All participants benefitted from the vocabulary pre- and post-tasks in different ways. The pre-tasks introduced the TWs, which likely ensured some level of familiarity with the terms. Similarly, both the vocabulary post-tasks and the vocabulary tests focused on form recall, which may have facilitated word form learning by providing participants with eight practice opportunities throughout the term, thereby reinforcing their memory of the TWs.

Moreover, since participants engaged in the same types of tasks and tests throughout the academic year, the test-retest effect –the improvement in performance due to the repetition of similar activities (Barenberg et al., 2021; van den Broek et al., 2013)– might have also contributed to the learning of TWs by both the CG and the EG.

In addition to the reasons outlined above, the lack of significant differences between the EG and the CG at some testing points could be attributed to the novelty of the experience for the EG. According to a background questionnaire completed by learners' relatives⁴, only 4% of EG participants watched OV TV on a weekly basis, and most were unfamiliar with this practice, despite watching L1 TV was their favorite leisure activity. Consequently, EG learners needed to adapt to processing a new type of audiovisual input, whereas CG participants were accustomed to focus-on-forms tasks, which are standard in the EFL classroom in the Catalan school context.

After some months of planned exposure to OV TV, significant differences between groups emerged, revealing a shift in trends. While most vocabulary learning continued occurring through explicit instruction, differences were observed in the

second and third terms, indicating that additional exposure to OV TV facilitated to a certain extent vocabulary learning. These findings align with previous research involving older learners (e.g., Montero Perez, 2020; Rodgers & Webb, 2020) and children with a relatively low level in the TL (Alexiou, 2015; Avello, 2023; Casulleras, 2023; Kuppens, 2010; Teng, 2019b; Teng & Cui, 2024).

As the academic year progressed, EG learners became more accustomed to watching OV TV, accumulating a total of eight hours and thirty-six minutes of additional input by the end of the intervention. Lexical gains in favor of the EG began to emerge at the end of the second term, suggesting that increased familiarity with subtitled TV series may have contributed to this improvement. This difference may be also linked to the fact that, after the first term, learners in the EG had become familiar with the dynamics of the intervention. As a result, they may have anticipated that the words presented in the pre- and post-tasks would be tested at the end of the term. In this context, they might have made a greater effort to commit TWs to memory and/or actively search for them in the episodes they were watching, thereby increasing learning opportunities. However, these gains were not consistent over time, and no significant differences in word meaning learning were observed at the end of the academic year, contrary to findings reported at more advanced levels (Gesa & Miralpeix, 2023). This inconsistency could be attributed to the challenge posed by the TV series for EG learners, who had a mean VS of 1,361 words. They likely encountered unfamiliar words that may have hindered their ability to fully benefit from the intervention, given that knowledge of the first 3,000 word families in English was needed for 95% lexical coverage, which is considered sufficient for video comprehension (Durbahn et al.,

2024). Consequently, EG learners may have prioritized word form learning, which was more manageable for them in this context.

Similarly, lexical gains were modest, likely due to the use of recall tests rather than recognition, as recall knowledge is generally more challenging to master than recognition knowledge (e.g., González-Fernández & Schmitt, 2020). Additionally, the stringent spelling-based scoring criteria for word form learning may have contributed to the relatively low vocabulary learning indices, despite the support provided by explicit teaching.

However, the limited lexical gains observed in the study are consistent with previous research on audiovisual input with children and adolescents testing recall knowledge. For instance, Pujadas and Muñoz (2019) found that eighth graders exposed to L1-subtitled TV series and pre-taught the TWs (similar to the EG in the present study) could recall 21.53% of word forms and 8.45% of word meanings. The results also aligned with those of Avello (2023), who found that primary school students in Chile could recall 17.64% of word forms after watching captioned TV series and completing construction-focused activities, a condition comparable to the EG in the present study. Likewise, Teng (2019b) reported that G6 learners could recall an average of 38.6% of TWs after exposure to 30 minutes of captioned videos. Greater learning was reported by Teng and Cui (2024), who found that G6 learners could recall 45.97% of word meanings under the full captioning condition. However, it should be noted that their participants were studying English in Hong Kong, where English is an official language and it is typically taught from kindergarten onwards. It should be noted that higher gains have also been reported with young learners in other studies (e.g., Alexiou, 2015; Casulleras, 2023; Chen et al., 2018; Koolstra & Beentjes, 1999; Neuman &

Koskinen, 1992; Suter et al., 2024; Teng, 2022b), although these studies typically measured form and meaning recognition rather than recall.

The amount of audiovisual input to which EG participants were exposed should also be considered: eight and a half hours of TV viewing over the entire academic year are relatively modest compared to the amount of input learners might receive in naturalistic settings or the average daily TV consumption in Spain, said to be 187 minutes in 2023 (Asociación para la Investigación de Medios de Comunicación, 2024). In this context, it could be argued that explicit teaching played a significant role, as the limited amount of audiovisual input meant that vocabulary recycling was restricted. Nevertheless, the results suggest that this practice might have a more substantial impact over time if there is extensive exposure to this type of input and repeated encounters with the target vocabulary are ensured (e.g., Uchihara et al., 2019). This is consistent with Webb's (2015) suggestion that the in-class viewing experience should aim at laying the groundwork for more extensive extramural viewing, encouraging learners to continue engaging with the TL independently.

5.2 The Role of Vocabulary Size and Language Aptitude in L2 Vocabulary Learning from Video Viewing and Teaching

Previous research with older learners (e.g., Montero Perez, 2020; Muñoz et al., 2023; Peters & Webb, 2018; Suárez & Gesa, 2019) has demonstrated a relationship between L2 VS and vocabulary learning from video viewing. The findings in this study provide evidence that a similar relationship exists among children with limited L2 processing abilities, consistent with the studies by Avello (2023) and Casulleras (2023). However, the present results contradict previous research with young learners that failed to find such a positive relationship (Teng & Cui, 2024). Overall, though, this close

relationship between VS and vocabulary learning is somewhat expected, as learners with a larger VS are generally more proficient and, therefore, more likely to benefit more from both explicit teaching and additional TV viewing (Chen et al., 2018; Gesa & Miralpeix, 2022).

Furthermore, learners with a larger VS may have found the task easier and more accessible: they likely recognized more vocabulary in the pre-tasks and had a higher level of coverage, which would enhance their comprehension and recognition of new words. A larger VS might have also facilitated language processing in the video, enabling learners to recognize more forms and follow the plot of the episodes more easily, thereby potentially enhancing the chances of word form learning. Regarding word meaning, learners with a larger VS could draw on their existing vocabulary knowledge during the pre-tasks to hypothesize about the meanings of unknown TWs. These hypotheses might have been confirmed or revised while watching the TV series, aided by contextual examples and images illustrating the meanings of the TWs.

Turning to language aptitude, it was found to mediate the number of word forms learned, though not the number of word meanings. The results for word form learning align with previous research on young learners, which has shown that language aptitude positively influences overall achievement in the FL (Kiss & Nikolov, 2005; Roehr-Brackin & Tellier, 2019). It is possible that higher aptitude assisted in learning the TWs during the pre-tasks. Additionally, aptitude may have also facilitated the efficient processing of audiovisual input and the integration of L1 subtitles with images and L2 audio. Consequently, participants with a higher aptitude could allocate more attentional resources to the vocabulary learning task, as processing the input may have been easier for them.

In contrast, language aptitude did not aid in learning TW meanings, possibly due to the high cognitive demands of the task. Establishing and retaining form-meaning links might have been too difficult for G6 participants, rendering higher language aptitude insufficient to overcome cognitive overload, as there was a significant amount of information presented rapidly and simultaneously in two different languages, one of which was partially unknown to these young learners. To learn word meanings, EG learners had to recognize the L2 word forms from the speech stream and assign the correct meaning based on the L1 subtitles and the visual support provided by the TV series. Similarly, during the pre-task phase, participants needed to get familiar with a large number of word forms and meanings in a short period of time (five words in approximately ten minutes), which could have exceeded their levels of language aptitude. This cognitive overload likely forced students to rely on other cognitive skills such as attention control, working memory, or long-term memory –skills not assessed by the MLAT-EC. This reliance on additional cognitive skills may also help to explain the low learning indices observed in the present study (Teng, 2023; Teng & Cui, 2024).

6. Pedagogical Implications

The present study is likely the first to assess the effects of additional TV viewing for vocabulary instruction among young learners. In a context where instruction time is very much limited, and teachers and practitioners must maximize its effectiveness, it is crucial to evaluate what supplementary activities can enhance traditional vocabulary teaching. This study has shown that supplementing traditional vocabulary instruction with OV TV viewing can be beneficial for vocabulary learning, particularly for word form learning. However, it is clear that the benefits of TV viewing were moderated by the influence of explicit teaching. This method could help young learners to make the

most of their FL instruction time, partly due to the scaffolding provided through pre-teaching. Moreover, significant differences between groups were found in the second and third terms, once the EG became familiar with OV TV viewing and the dynamics of the intervention. This suggests that audiovisual input should be integrated into the school curriculum as a long-term practice rather than used sporadically. Consistent implementation of this approach over time could maximize lexical gains, highlighting its potential for enhancing vocabulary acquisition.

Lastly, the study demonstrated that VS and to a much lesser extent language aptitude mediate vocabulary learning from explicit teaching and additional exposure to extensive viewing. This finding has important implications for the language classroom. First, careful selection of audiovisual materials is necessary for extensive viewing practices, as the coverage learners have of the OV input will either enhance or hinder lexical learning. Regarding aptitude, the results indicate that it is not a highly significant factor in determining the effectiveness of additional OV input exposure. Therefore, it can be concluded that all learners could benefit from such practice, regardless of their language aptitude, as long as they are granted the appropriate scaffolding through vocabulary learning activities.

7. Conclusions and Further Research

The present study explored the use of extensive viewing in addition to intentional learning to enhance vocabulary acquisition among primary school students. Despite previous research on the effects of subtitled video viewing for L2 learning, there exists a need for more classroom-based research with young learners (Montero Perez, 2022; Reynolds et al., 2022). This study addressed this gap, and its results demonstrated the potential of combining teaching and viewing for vocabulary

development in the primary school classroom. It also introduces a novel methodological approach to boost vocabulary learning by integrating an informal activity, such as TV viewing, into the school curriculum (Teng, 2025).

The study also opens several lines of enquiry. First, significant improvement in word form learning was observed with L1 subtitles, suggesting the potential benefit of other types of textual support (e.g., enhanced or glossed captions) for learners at lower levels. Additionally, participants in the EG watched the episodes once only. Future research should investigate the benefits of repeated viewing for young learners. While this study focused on vocabulary learning, it is important to explore potential gains in other areas, such as viewing comprehension, pronunciation development, and reading efficacy. Furthermore, this paper examined the influence of L2 VS and language aptitude on learning from extensive viewing. However, other variables, such as working memory, and different word factors (e.g., concreteness, frequency of occurrence, imagery) could also affect lexical development and should be investigated in future research. Lastly, this study primarily used quantitative measures of vocabulary learning, but it did not investigate learners' actual thought processes. Therefore, it would be valuable to triangulate these measures with qualitative data (e.g., stimulated recall) to gain a deeper understanding of how vocabulary learning from OV TV viewing occurs (Montero Perez et al., 2025). These suggestions, together with the findings of the present study, will contribute to advancing our understanding of FL learning and teaching among young learners.

Notes

¹ All pre- and post-tests, as well as pre- and post-tasks, have been uploaded to the IRIS repository.

² They included multiple choice and true / false questions, and an event-ordering exercise. However, these comprehension tests will not be analysed in this paper since they are beyond the scope of the present study.

³ For the descriptive statistics of absolute gains and raw scores on pre- and post-tests, please refer to Tables B1 and B2 in Appendix B.

⁴ This background questionnaire was completed at the beginning of the teaching intervention and included questions about participants' age, L1s, socioeconomic status, and their extramural exposure to the TL (e.g., frequency of TV viewing, reading or gaming, attendance at private classes, and experiences abroad or at summer camps).

References

- Ahrabi Fakhr, M., Borzabadi Farahani, D., & Khomeijani Farahani, A. A. (2021). Incidental vocabulary learning and retention from audiovisual input and factors affecting them. *English Teaching & Learning*, 45, 167–188.
<https://doi.org/10.1007/s42321-020-00066-y>
- Alexiou, T. (2015). Vocabulary uptake from *Peppa Pig*: A case study of preschool EFL learners in Greece. In C. Gitsaki & T. Alexiou (Eds.), *Current issues in second / foreign language teaching and teacher development: Research and practice* (pp. 285–301). Cambridge Scholars Publishing.
- Arifani, Y. (2020). Cartoon video-assisted learning: An investigation into the acquisition of EFL children's incidental vocabulary. *Computer-Assisted Language Learning Electronic Journal*, 21(2), 17–31.
- Asociación para la Investigación de Medios de Comunicación [AIMC]. (2024). *Marco general de los medios en España 2024* [General media survey in Spain 2024].

- AIMC. https://www.aimc.es/aimc-content/uploads/2024/01/Marco_General_Medios_2024.pdf. Accessed 11 November 2024
- Ausubel, D. P. (1968). *Educational psychology: A cognitive view*. Holt, Rinehart & Winston.
- Avello, D. I. (2023). *L2 learning from captioned-video viewing in primary school students* (Unpublished doctoral dissertation). Universitat de Barcelona, Barcelona, Spain.
- Babaei, S., & Izadpanah, S. (2019). Comparing the effects of different advance organizers on EFL learners' listening comprehension: Key vocabularies, previewing comprehension questions, and multimedia annotations. *Cogent Education*, 6(1), 1705666. <https://doi.org/10.1080/2331186X.2019.1705666>
- Barcroft, J. (2015). *Lexical input processing and vocabulary learning*. John Benjamins.
- Barenberg, J., Berse, T., Reimann, L., & Dutke, S. (2021). Testing and transfer: Retrieval practice effects across test formats in English vocabulary learning in school. *Applied Cognitive Psychology*, 35(3), 700–710. <https://doi.org/10.1002/acp.3796>
- Cadierno, T., Hansen, M., Lauridsen, J. T., Eskildsen, S. W., Fenyvesi, K., Jensen, S. H., & aus der Wieschen, M. V. (2020). Does younger mean better? Age of onset, learning rate and short-term L2 proficiency in young Danish learners of English. *Vigo International Journal of Applied Linguistics*, 17, 57–86. <https://doi.org/10.35869/vial.v0i17.1465>
- Caimi, A. (2006). Audiovisual translation and language learning: The promotion of intralingual subtitles. *The Journal of Specialised Translation*, 6, 85–98.

- Carroll, J. B., & Sapon, S. (1967). *Modern Language Aptitude Test-Elementary*. Psychological Corporation.
- Casulleras, M. (2023). *Language learning from subtitled TV series in the primary EFL classroom* (Unpublished doctoral dissertation). Universitat de Barcelona, Barcelona, Spain.
- Cenoz, J. (2002). Age differences in foreign language learning. *ITL - International Journal of Applied Linguistics*, 135(1), 125–142. <https://doi.org/10.1075/itl.135-136.06cen>
- Chen, Y.-R., Liu, Y.-T., & Todd, A. G. (2018). Transient but effective? Captioning and adolescent EFL learners' spoken vocabulary acquisition. *English Teaching & Learning*, 42(1), 25–56. <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 / La Revue TESL du Canada*, 14(1), 61–65. <https://doi.org/10.18806/tesl.v14i1.678>
- Chung, J.-M. (2002). The effects of using two advance organizers with video texts for the teaching of listening in English. *Foreign Language Annals*, 35(2), 231–241. <https://doi.org/10.1111/j.1944-9720.2002.tb03157.x>
- Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Lawrence Erlbaum Associates.
- Davies, M. (2009). The 385+ million word *Corpus of Contemporary American English* (1990–2008+): Design, architecture, and linguistic insights. *International Journal of Corpus Linguistics*, 14(2), 159–190. <https://doi.org/10.1075/ijcl.14.2.02dav>

- De Wilde, V., Brysbaert, M., & Eyckmans, J. (2020). Learning English through out-of-school exposure. Which levels of language proficiency are attained and which types of input are important? *Bilingualism: Language and Cognition*, 23(1), 171–185. <http://doi.org/10.1017/S1366728918001062>
- De Wilde, V., Brysbaert, M., & Eyckmans, J. (2022). Formal versus informal L2 learning. How do individual differences and word-related variables influence French and English L2 vocabulary learning in Dutch-speaking children? *Studies in Second Language Acquisition*, 44(1), 87–111. <http://doi.org/10.1017/S0272263121000097>
- Díaz Cintas, J. (2012). Subtitling. Theory, practice and research. In C. Millán & F. Bartrina (Eds.), *The Routledge handbook of translation studies* (pp. 273–287). Routledge.
- Durbahn, M., Rodgers, M., Macis, M., & Peters, E. (2024). Lexical coverage in L1 and L2 viewing comprehension. *Studies in Second Language Acquisition, FirstView*, 1–24. <https://doi.org/10.1017/S0272263124000391>
- d'Ydewalle, G., & Pavakanun, U. (1997). Could enjoying a movie lead to language acquisition? In P. Winterhoff-Spurk & T. H. A. van der Voort (Eds.), *New horizons in media psychology: Research cooperation and projects in Europe* (pp. 145–155). Springer Fachmedien.
- d'Ydewalle, G., & Van de Poel, M. (1999). Incidental foreign-language acquisition by children watching subtitled television programs. *Journal of Psycholinguistic Research*, 28(3), 227–244.

- Elkhafaifi, H. (2005). The effect of prelistening activities on listening comprehension in Arabic learners. *Foreign Language Annals*, 38(4), 505–513.
<https://doi.org/10.1111/j.1944-9720.2005.tb02517.x>
- Ellis N. C. (2002). Frequency effects in language processing. *Studies in Second Language Acquisition*, 24(2), 143–188.
<https://doi.org/10.1017/S0272263102002024>
- Gathercole, S. E., & Baddeley, A. D. (1993). *Working memory and language*. Lawrence Erlbaum Associates, Inc.
- Gesa, F., & Miralpeix, I. (2022). Effects of watching subtitled TV series on foreign language vocabulary learning: Does learners' proficiency level matter? In C. Lütge (Ed.), *Foreign language learning in the digital age. Theory and pedagogy for developing literacies* (pp. 159–173). Routledge.
- Gesa, F., & Miralpeix, I. (2023). Extensive viewing as additional input for foreign language vocabulary learning: A longitudinal study in secondary school. *Language Teaching Research*, OnlineFirst.
<https://doi.org/10.1177/13621688231169451>
- González-Fernández, B., & Schmitt, N. (2020). Word knowledge: Exploring the relationships and order of acquisition of vocabulary knowledge components. *Applied Linguistics*, 41(4), 481–505. <https://doi.org/10.1093/applin/amy057>
- Greenwald, T. J., Murrieta, P., Cheung, V., & Montanio, B. (Producers). (2007). *Wizards of Waverly Place* [Television series]. Disney Channel.
- Hanley, J. E. B., Herron, C. A., & Cole, S. P. (1995). Using video as an advance organizer to a written passage in the FLES classroom. *The Modern Language Journal*, 79(1), 57–66. <https://doi.org/10.1111/j.1540-4781.1995.tb05416.x>

- Hariffin, A., & Said, N. E. M. (2019). The effects of captioned videos on primary ESL learners' vocabulary acquisition in a Malaysian rural setting. *International Journal of Humanities, Philosophy, and Language*, 2(7), 23–37.
- Horst, M., Cobb, T., & Meara, P. M. (1998). Beyond a clockwork orange: Acquiring second language vocabulary through reading. *Reading in a Foreign Language*, 11(2), 207–223.
- Jafari, K., & Hashim, F. (2012). The effects of using advance organizers on improving EFL learners' listening comprehension: A mixed method study. *System*, 40(2), 270–281. <http://dx.doi.org/10.1016%2Fj.system.2012.04.009>
- Jiang, X., & Grabe, W. (2007). Graphic organizers in reading instruction: Research findings and issues. *Reading in a Foreign Language*, 19(1), 34–55.
<https://doi.org/10125/66616>
- Kallis, D., Dreayer, I., Eells, P., & Geoghan, J. (Producers). (2005). *The Suite Life of Zack and Cody* [Television series]. Disney Channel.
- Kiss, C., & Nikolov, M. (2005). Developing, piloting, and validating an instrument to measure young learners' aptitude. *Language Learning*, 55(1), 99–150.
<https://doi.org/10.1111/j.0023-8333.2005.00291.x>
- Koolstra, C. M., & Beentjes, J. W. J. (1999). Children's vocabulary acquisition in a foreign language through watching subtitled television programs at home. *Educational Technology Research and Development*, 47(1), 51–60.
<https://doi.org/10.1007/BF02299476>
- Kuppens, A. H. (2010). Incidental foreign language acquisition from media exposure. *Learning, Media and Technology*, 35(1), 65–85.
<https://doi.org/10.1080/17439880903561876>

- Laufer, B. (2006). Comparing focus on form and focus on formS in second-language vocabulary learning. *The Canadian Modern Language Review / La revue canadienne des langues vivantes*, 63(1), 149–166.
<https://doi.org/10.3138/cmlr.63.1.149>
- Laufer, B., & Goldstein, Z. (2004). Testing vocabulary knowledge: Size, strength, and computer adaptiveness. *Language Learning*, 54(3), 399–436.
<https://doi.org/10.1111/j.0023-8333.2004.00260.x>
- Laufer, B., & Rozovski-Roitblat, B. (2015). Retention of new words: Quantity of encounters, quality of task, and degree of knowledge. *Language Teaching Research*, 19(6), 687–711. <https://doi.org/10.1177/1362168814559797>
- Lee, J. H., & Lee, H. (2022). Teachers’ verbal lexical explanation for second language vocabulary learning: A meta-analysis. *Language Learning*, 72(2), 576–612.
<https://doi.org/10.1111/lang.12493>
- Li, C.-H., Wu, M.-H., & Lin, W.-L. (2019). The use of a “Think-Pair-Share” brainstorming advance organizer to prepare learners to listen in the L2 classroom. *International Journal of Listening*, 33(2), 114–127.
<https://doi.org/10.1080/10904018.2017.1394193>
- Lindstromberg, S. (2020). Intentional L2 vocabulary learning. In S. Webb (Ed.), *The Routledge handbook of vocabulary studies* (pp. 240–254). Routledge.
- Linebarger, D. L., Moses, A., Garrity Liebeskind, K., & McMnamin, K. (2013). Learning vocabulary from television: Does onscreen print have a role? *Journal of Educational Psychology*, 105(3), 609–621. <https://doi.org/10.1037/a0032582>

- Luquin, M., & García Mayo, M. d. P. (2021). Exploring the use of models as a written corrective feedback technique among EFL children. *System*, 98, 102465.
<https://doi.org/10.1016/j.system.2021.102465>
- Meara, P. M. (2005). X_Lex: the Swansea vocabulary levels test (Version 2.05) [Computer software]. Centre for Applied Language Studies, University of Wales, the United Kingdom.
- Miralpeix, I. (2008). *The influence of age on vocabulary acquisition in English as a foreign language* (Unpublished doctoral dissertation). Universitat de Barcelona, Barcelona, Spain.
- Montero Perez, M. (2020). Incidental vocabulary learning through video viewing: The role of vocabulary knowledge and working memory. *Studies in Second Language Acquisition*, 42(4), 749–773.
<http://doi.org/10.1017/S0272263119000706>
- Montero Perez, M. (2022). Second or foreign language learning through watching audio-visual input and the role of on-screen text. *Language Teaching*, 55(2), 163–192. <http://doi.org/10.1017/S0261444821000501>
- 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.
<http://dx.doi.org/10125/44357>
- Montero Perez, M., Peters, E., & Desmet, P. (2015). Enhancing vocabulary learning through captioned video: An eye-tracking study. *The Modern Language Journal*, 99(2), 308–328. <https://doi.org/10.1111/modl.12215>

- Montero Perez, M., Peters, E., & Desmet, P. (2018). Vocabulary learning through viewing video: The effect of two enhancement techniques. *Computer Assisted Language Learning*, 31(1-2), 1–26.
<https://doi.org/10.1080/09588221.2017.1375960>
- Montero Perez, M., Sydorenko, T., & Huntley, L. (2025). Incidental vocabulary learning from watching audiovisual input. In M. F. Teng & B. L. Reynolds (Eds.), *Researching incidental vocabulary learning in a second language* (pp. 35–50). Routledge.
- Montero Perez, M., Van Den Noortgate, W., & Desmet, P. (2013). Captioned video for L2 listening and vocabulary learning: A meta-analysis. *System*, 41(3), 720–739.
<https://doi.org/10.1016/j.system.2013.07.013>
- Muñoz, C. (2008). Symmetries and asymmetries of age effects in naturalistic and instructed L2 learning. *Applied Linguistics*, 29(4), 578–596.
<https://doi.org/10.1093/applin/amm056>
- Muñoz, C. (2020). Boys like games and girls like movies. Age and gender differences in out-of-school contact with English. *Revista Española de Lingüística Aplicada / Spanish Journal of Applied Linguistics*, 33(1), 171–201.
<https://doi.org/10.1075/resla.18042.mun>
- Muñoz, C., Pattenmore, A., & Avello, D. (2022). Exploring repeated captioning viewing as a way to promote vocabulary learning: Time lag between repetitions and learner factors. *Computer Assisted Language Learning*, Latest Articles.
- Muñoz, C., Pujadas, G., & Pattenmore, 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., & Beglar, D. (2007). A vocabulary size test. *The Language Teacher*, 31(7), 9–13.
- Nation, P., & Webb, S. (2011). *Researching and analyzing vocabulary*. Heinle Cengage Learning.
- Neuman, S. B., & Koskinen, P. (1992). Captioned television as comprehensible input: Effects of incidental word learning from context for language minority students. *Reading Research Quarterly*, 27(1), 94–106. <https://doi.org/10.2307/747835>
- Patterson, J. L. (2002). Relationships of expressive vocabulary to frequency of reading and television experience among bilingual toddlers. *Applied Psycholinguistics*, 23, 493–508. <https://doi.org/10.1017/S0142716402004010>
- Peters, E., Heynen, E., & Puimège, E. (2016). Learning vocabulary through audiovisual input: The differential effect of L1 subtitles and captions. *System*, 63, 134–148. <https://doi.org/10.1016/j.system.2016.10.002>
- Peters, E., & Webb, S. (2018). Incidental vocabulary acquisition through viewing L2 television and factors that affect learning. *Studies in Second Language Acquisition*, 40(3), 551–577. <https://doi.org/10.1017/S0272263117000407>
- Plonsky, L., & Oswald, F. L. (2014). How big is “big”? Interpreting effect sizes in L2 research. *Language Learning*, 64(4), 878–912. <https://doi.org/10.1111/lang.12079>
- Puimège, E., & Peters, E. (2019). Learners’ English vocabulary knowledge prior to formal instruction: The role of learner-related and word-related variables. *Language Learning*, 69(4), 943–977. <https://doi.org/10.1111/lang.12364>
- 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. (2023). Measuring the visual in audio-visual input. The effects of imagery in vocabulary learning through TV viewing. *ITL -*

International Journal of Applied Linguistics, 174(2), 263–290.

<https://doi.org/10.1075/itl.22019.puj>

Reynolds, B. L., Cui, Y., Kao, C.-W., & Thomas, N. (2022). Vocabulary acquisition through viewing captioned and subtitled video: A scoping review and meta-

analysis. *Systems*, 10(5), 133. <https://doi.org/10.3390/systems10050133>

Rice, M. L., Huston, A. C., Truglio, R., & Wright, J. (1990). Words from “Sesame Street”: Learning vocabulary while viewing. *Developmental Psychology*, 26(3),

421–428. <https://doi.org/10.1037/0012-1649.26.3.421>

Richardson, J. T. E. (2011). Eta squared and partial eta squared as measures of effect size in educational research. *Educational Research Review*, 6(2), 135–147.

<https://doi.org/10.1016/j.edurev.2010.12.001>

Rinehart, S. D., & Welker, W. A. (1992). Effects of advance organizers on level and time of text recall. *Reading Research and Instruction*, 32(1), 77–86.

<https://doi.org/10.1080/19388079209558107>

Rispa, J. (2021). *The effects of language aptitude on L2 development in young learners*.

(Unpublished bachelor’s thesis). Universitat de Lleida, Lleida, Spain.

Rodgers, M. P. H. (2016). Extensive listening and viewing: The benefits of audiobooks and television. *The European Journal of Applied Linguistics and TEFL*, 5(2),

43–57.

- Rodgers, M. P. H., & Webb, S. (2020). Incidental vocabulary learning through viewing television. *ITL - International Journal of Applied Linguistics*, 171(2), 191–220. <https://doi.org/10.1075/itl.18034.rod>
- Roehr-Brackin, K., & Tellier, A. (2019). The role of language-analytic ability in children's instructed second language learning. *Studies in Second Language Acquisition*, 41(5), 1111–1131. <https://doi.org/10.1017/S0272263119000214>
- Schmidt, R. (2012). Attention, awareness, and individual differences in language learning. In W. M. Chan, K. N. Chin, S. K. Bhatt & I. Walker (Eds.), *Perspectives on individual characteristics and foreign language education* (pp. 27–50). De Gruyter Mouton.
- Schmitt, N. (2008). Instructed second language vocabulary learning. *Language Teaching Research*, 12(3), 329–363. <https://doi.org/10.1177/1362168808089921>
- Schmitt, N., & Schmitt, D. (2020). *Vocabulary in language teaching* (2nd ed.). Cambridge University Press.
- Severson, K. J. (2017). *Interactive videos as a vocabulary pre-teaching tool in middle school science*. (Unpublished master's thesis). Hamline University, Saint Paul, the United States.
- Skehan, P. (1991). Individual differences in second language learning. *Studies in Second Language Acquisition*, 13(2), 275–298. <http://doi.org/10.1017/S0272263100009979>
- Suárez, M. M. (2010). *Language aptitude in young learners: The Elementary Modern Language Aptitude Test in Spanish and Catalan* (Unpublished doctoral dissertation). Universitat de Barcelona, Barcelona, Spain.

- Suárez, M. M., & Gesa, F. (2019). Learning vocabulary with the support of sustained exposure to captioned video: Do proficiency and aptitude make a difference? *The Language Learning Journal*, 47(4), 497–517.
<https://doi.org/10.1080/09571736.2019.1617768>
- Suter, L., Berthele, R., & Thomas, A. (2024). L'apprentissage du vocabulaire en français L2 avec Netflix: Effet des activités de pré-/post-visionnage [Learning French L2 vocabulary with Netflix: Effect of pre-/post-viewing activities]. *Journal of French Language Studies*, First View.
<https://doi.org/10.1017/S0959269523000273>
- Tavakol, M., & Dennick, R. (2011). Making sense of Cronbach's alpha. *International Journal of Medical Education*, 2, 53–55. <https://doi.org/10.5116/ijme.4dfb.8dfd>
- Teng, M. F. (2019a). The effects of video caption types and advance organizers on incidental L2 collocation learning. *Computers & Education*, 142, 1036555.
<https://doi.org/10.1016/j.compedu.2019.103655>
- Teng, M. F. (2019b). Incidental vocabulary learning for primary school students: The effects of L2 caption type and word exposure frequency. *The Australian Educational Researcher*, 46, 113–136. <https://doi.org/10.1007/s13384-018-0279-6>
- Teng, M. F. (2022a). Incidental L2 vocabulary learning from viewing captioned videos: Effects of learner-related factors. *System*, 105, 102736.
<https://doi.org/10.1016/j.system.2022.102736>
- Teng, M. F. (2022b). Vocabulary learning through videos: Captions, advance-organizer strategy, and their combination. *Computer Assisted Language Learning*, 35(3), 518–550. <https://doi.org/10.1080/09588221.2020.1720253>

- Teng, M. F. (2023). Effectiveness of captioned videos for incidental vocabulary learning and retention: The role of working memory. *Computer Assisted Language Learning*, Latest Articles. <https://doi.org/10.1080/09588221.2023.2173613>
- Teng, M. F. (2025). Introduction to *Researching incidental vocabulary learning in a second language*. In M. F. Teng & B. L. Reynolds (Eds.), *Researching incidental vocabulary learning in a second language* (pp. 1–15). Routledge.
- Teng, M. F., & Cui, Y. (2024). Comparing incidental learning of single words and collocations from different captioning conditions: The role of vocabulary knowledge and working memory. *Journal of Computer Assisted Learning*, 40(3), 973–989. <https://doi.org/10.1111/jcal.12910>
- Uchihara, T., Webb, S., & Yanagisawa, A. (2019). The effects of repetition on incidental vocabulary learning: A meta-analysis of correlational studies. *Language Learning*, 69(3), 559–599. <https://doi.org/10.1111/lang.12343>
- van den Broek, G. S. E., Takashima, A., Segers, E., Fernández, G., & Verhoeven, L. (2013). Neural correlates of testing effects in vocabulary learning. *NeuroImage*, 78, 94–102. <https://doi.org/10.1016/j.neuroimage.2013.03.071>
- Vanderplank, R. (2016). *Captioned media in foreign language learning and teaching: Subtitles for the deaf and hard-of-hearing as tools for language learning*. Palgrave Macmillan.
- Vu, D. V., Noreillie, A.-S., & Peters, E. (2023). Incidental collocation learning from reading-while-listening and captioned TV viewing and predictors of learning gains. *Language Teaching Research*, Online First. <https://doi.org/10.1177/13621688221151048>

- Walker, I., & Hulme, C. (1999). Concrete words are easier to recall than abstract words: Evidence for a semantic contribution to short-term serial recall. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 25(5), 1256–1271. <https://doi.org/10.1037/0278-7393.25.5.1256>
- Webb, S. (2007). The effects of repetition on vocabulary knowledge. *Applied Linguistics*, 28(1), 46–65. <https://doi.org/10.1093/applin/aml048>
- 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. (2020). Incidental vocabulary learning. In S. Webb (Ed.), *The Routledge handbook of vocabulary studies* (pp. 225–239). Routledge.
- Webb, S., & Nation, P. (2017). *How vocabulary is learned*. Oxford University Press.
- Webb, S., & Rodgers, M. P. H. (2009). Vocabulary demands of television programs. *Language Learning*, 59(2), 335–366. <https://doi.org/10.1111/j.1467-9922.2009.00509.x>
- Webb, S., Sasao, Y., & Ballance, O. (2017). The updated Vocabulary Levels Test: Developing and validating two new forms of the VLT. *ITL - International Journal of Applied Linguistics*, 168(1), 33–69. <https://doi.org/10.1075/itl.168.1.02web>
- Wilberschied, L., & Berman, P. M. (2004). Effect of using photos from authentic video as advance organizers on listening comprehension in an FLES Chinese class.

Foreign Language Annals, 37(4), 534–540. <https://doi.org/10.1111/j.1944-9720.2004.tb02420.x>

Wright, J. C., Huston, A. C., Murphy, K. C., Saint Peters, M., Piñon, M., Scantlin, R., & Kotler, J. (2001). The relations of early television viewing to school readiness and vocabulary of children from low-income families: The Early Window Project. *Child Development*, 72(5), 1347–1366. <https://doi.org/10.1111/1467-8624.t01-1-00352>

Yang, H.-Y. (2014). The effects of advance organizers and subtitles on EFL learners' listening comprehension skills. *CALICO Journal*, 31(3), 345–373. <https://doi.org/10.11139/cj.31.3.345-373>

Zhang, X., Liu, J., & Ai, H., (2019). Pseudowords and guessing in the Yes/No format vocabulary test. *Language Testing*, 37(1), 6–30. <https://doi.org/10.1177/0265532219862265>

Appendix A

List of TWs on Which Participants Were Tested, and Their Frequency of Occurrence in the Episode and in Larger Corpora (Davies, 2009)

Term 1			Term 2			Term 3		
TW	Frequency		TW	Frequency		TW	Frequency	
	Episode (<i>N</i> times)	Band		Episode (<i>N</i> times)	Band		Episode (<i>N</i> times)	Band
Admirer	9	2K	Beneath	4	1K	Ant	5	5K
Ankle	3	4K	Birth certificate	3	4K	Apron	3	6K
To apologize	2	3K	To blot	2	6K	Backpack	2	Off-list
Awesome	2	6K	Bossy	5	2K	To bark	2	2K
Badge	2	5K	To care	4	1K	Basement	5	4K
Ballroom	3	Off-list	Chickenpox	2	Off-list	Brick	11	2K
Beauty mark	2	1K	Cookie	9	7K	Cage	3	2K
Burden	4	3K	Crook	7	5K	Carpet	12	2K
Concierge	3	12K	Diaper	3	10K	Charm	6	2K
Contest	7	3K	Dungeon	3	10K	To cheat	6	2K
Counter	2	2K	To eavesdrop	2	9K	Commitment	3	2K
Cute	5	6K	To fast forward	5	1K	To curl	5	2K
Date	5	1K	To fix	6	1K	Dandruff	2	13K
To ditch	5	5K	To flick	3	4K	Drill	5	3K
Dweeb	4	20K	To flunk	2	13K	Elf	28	5K
Goat	2	4K	Footprint	3	2K	Essay	3	3K
Grounded	5	1K	Fox	4	2K	To fail	14	2K
To hurt	6	1K	Frizzy	3	12K	To fire	14	1K
Lame	5	7K	Gearshift	2	Off-list	Fur	4	2K
Lifeguard	4	Off-list	To gossip	9	4K	Freewheeling	3	Off-list
Lipstick	2	Off-list	Jewel	13	5K	Goth	5	Off-list
Lobby	6	3K	To lock	4	1K	To hire	5	2K
To owe	2	2K	Loot	3	5K	Jail	3	3K
Pageant	11	7K	Lounge	5	10K	Lair	3	10K
Prize	2	2K	To mix up	3	2K	Owner	9	1K
Prom	23	9K	Nervous wreck	2	4K	Partner	4	2K
Review	4	3K	Outfit	3	4K	To pop	5	1K
Slumber party	2	9K	Peanut	6	6K	Potion	17	10K
Step	4	1K	Pie	4	2K	Purse	16	5K
Sweet	4	1K	Pillow	2	4K	Rightful	4	1K
Terrific	2	5K	To pledge	3	4K	To schmooze	5	15K
Thirsty	3	1K	Recess	4	6K	Sneakers	4	5K
Tux	3	Off-list	To rehearse	4	4K	Spell	8	2K
Vase	4	5K	Smart	5	2K	Summit	12	3K
Vent	6	5K	Spy gear	4	2K	Unsightly	5	1K
Wedding	9	1K	To stick	6	1K	Waitress	5	4K
Weenies	4	17K	To sweat	4	3K	Wallpaper	9	Off-list
To work out	3	1K	Trap	4	2K	Wand	16	9K
Wrestler	3	5K	Twin	7	2K	Wizard	13	5K
To yell	4	2K	Whistle	5	2K	Zit	12	18K

Appendix B

Table B1

Descriptive Statistics for Absolute Gains for Form and Meaning

Group	Form*			Meaning*		
	<i>M</i>	<i>SD</i>	95% CI	<i>M</i>	<i>SD</i>	95% CI
Term 1						
EG (<i>n</i> =22)	7.45	3.74	[5.80, 9.11]	3.95	3.15	[2.56, 5.35]
CG (<i>n</i> =18)	6.50	3.28	[4.87, 8.13]	3.89	3.64	[2.08, 5.70]
Term 2						
EG (<i>n</i> =23)	6.39	3.59	[4.84, 7.94]	4.30	2.38	[3.27, 5.33]
CG (<i>n</i> =17)	4.53	2.50	[3.24, 5.82]	3.41	3.36	[1.69, 5.14]
Term 3						
EG (<i>n</i> =22)	6.86	3.56	[5.28, 8.44]	4.32	2.51	[3.20, 5.43]
CG (<i>n</i> =18)	4.33	2.89	[2.90, 5.77]	3.11	2.25	[1.99, 4.23]

*The maximum score was 40.

Table B2*Raw Scores for Form and Meaning on the Pre- and Post-Tests*

Group	Test	Form*			Meaning*		
		<i>M</i>	<i>SD</i>	95% CI	<i>M</i>	<i>SD</i>	95% CI
Term 1							
EG (<i>n</i> =22)	Pre-	4.08	3.01	[2.84, 5.32]	1.32	1.82	[.57, 2.07]
	Post-	10.20	5.64	[7.87, 12.53]	5.16	3.82	[3.59, 6.73]
CG (<i>n</i> =18)	Pre-	3.05	3.03	[1.67, 4.42]	1.57	2.40	[.48, 2.66]
	Post-	8.57	5.70	[5.98, 11.16]	4.90	5.15	[2.56, 7.25]
Term 2							
EG (<i>n</i> =23)	Pre-	4.80	1.89	[4.02, 5.58]	1.36	1.22	[.86, 1.86]
	Post-	9.80	4.57	[7.91, 11.69]	7	3.54	[5.47, 8.53]
CG (<i>n</i> =17)	Pre-	3.39	2.06	[2.50, 4.28]	1.65	1.50	[1.01, 2.30]
	Post-	5.40	2.92	[4.20, 6.60]	4.04	4.19	[2.23, 5.86]
Term 3							
EG (<i>n</i> =22)	Pre-	6	3.32	[4.60, 7.40]	2.79	1.53	[2.14, 3.44]
	Post-	11.54	5.85	[9.07, 14.01]	6.92	3.56	[5.41, 8.42]
CG (<i>n</i> =18)	Pre-	3.46	2.39	[2.45, 4.47]	1.79	1.53	[1.14, 2.44]
	Post-	6.83	3.92	[5.18, 8.49]	4.71	2.90	[3.49, 5.93]

*The maximum score was 40.