

Article

# The Development of Receptive Language Skills from Captioned Video Viewing in Primary School EFL Learners

Daniela Avello \*  and Carmen Muñoz

Department of Modern Languages and Literatures and English Studies, University of Barcelona,  
08007 Barcelona, Spain; munoz@ub.edu

\* Correspondence: diag08041@gmail.com

**Abstract:** The evidence obtained to date supports the use of captioned videos for L2 learning purposes, such as vocabulary acquisition and the development of L2 listening skills. However, little research has been conducted with primary school learners, and even less so on the extent to which L2-captioned videos foster the development of L2 reading skills. Thus, the present investigation aimed to determine the extent to which five groups of primary school EFL learners from Chile ( $n = 96$ , 9–11 years old, years 4 and 5) benefited from their viewing experience (11 captioned videos) as regards the development of L2 listening skills and L2 reading efficacy (measured at pretest, posttest, and delayed-posttest). In addition, we assessed the influence of L1- and L2-related factors on learners' performance over time (L1 and L2 reading efficacy, L2 vocabulary knowledge, and L2 listening skills). Overall, the results revealed that the treatment led to significant gains in English listening skills and reading efficacy in fourth and fifth graders. However, learners' performance was also found to be predicted by language-related factors, especially L2 vocabulary knowledge. On the whole, the findings of this investigation support the use of age-appropriate captioned videos at primary school to increase children's exposure to the target language and enhance the development of receptive language skills.

**Keywords:** audiovisual input; young learners; foreign language learning; reading skills; listening skills; individual differences



**Citation:** Avello, D.; Muñoz, C. The Development of Receptive Language Skills from Captioned Video Viewing in Primary School EFL Learners. *Educ. Sci.* **2023**, *13*, 479. <https://doi.org/10.3390/educsci13050479>

Academic Editor: Martin Howard

Received: 10 March 2023

Revised: 4 May 2023

Accepted: 4 May 2023

Published: 9 May 2023



**Copyright:** © 2023 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

## 1. Introduction

The literature has increasingly provided evidence of the positive effects of audiovisual input on L2 learning [1,2]. Nonetheless, most of the investigations have been conducted with teenagers and adult L2 learners [3]. While various studies on out-of-school learning have demonstrated that primary school students do benefit from their exposure to audiovisual input, it is important to bear in mind that the extent to which learners watch TV in a foreign language from an early age depends on contextual factors and national policies [4–6]. In the case of input-limited contexts, where dubbing is usually the norm, learners' exposure to TV series and movies in a foreign language is very limited [4,7] and restricted to their access to video-sharing websites, pay-per-view TV channels, and streaming platforms. However, viewers may still control the settings and choose to watch the videos in their L1. In fact, the evidence has indicated that children and families are not necessarily aware of the potential benefits of audiovisual input and text support for L2 learning [4]. As a result, the implementation of a principled approach to viewing in the classroom, which develops families' awareness of the actual benefits and the strategies that may be used to improve the experience, might be key to encouraging L2 learners to do this activity at home and increase their exposure to the L2 [8].

There is little evidence on the effects of extensive viewing on the development of receptive L2 skills (e.g., [5]). Overall, the few existing experimental studies that have examined the effects of captioned video viewing on the development of listening skills

by means of fill-in-the-blanks and shadowing tasks have demonstrated that the use of onscreen text enhances speech perception, namely bottom-up processing, after a short intervention [9–11]. Still, it is uncertain whether these results could be replicated with young learners by means of short interventions because, in foreign language contexts, primary school students—whose cognitive skills are still developing—have been found to be less efficient learners than teenagers and adults [12]. As for the development of reading skills as a result of captioned video viewing, most of the studies that have explicitly focused on this issue have been conducted in L1 contexts [13–15]. Therefore, further research is required to determine whether the findings emerging from L1 settings may also be obtained in foreign language contexts. Given these gaps in the literature, the present investigation attempted to determine the extent to which extensive captioned video viewing fostered the development of receptive L2 skills in primary school learners and to explore the role of language-related factors (L1 and L2) on the outcomes. On the whole, the results of this investigation may strengthen the implementation of language learning programs with primary school students in input-limited foreign language learning contexts to enhance the development of receptive language skills [16–18]. These young learners, who are still developing their L1 reading skills, need to increase their exposure to the target language by means of activities that are suited to their characteristics.

### *1.1. Captions and L2 Listening Skills Development*

The literature suggests that listening comprehension tasks may be quite challenging for L2 learners, particularly at lower proficiency levels, because specific information and ideas must be extracted at a speed that listeners cannot control [19,20]. Existing investigations on captioned video viewing have shown evidence of the positive effects of onscreen text on viewing comprehension, which may be associated with speech segmentation and learners' capacity to identify word boundaries in the stream of speech [10]. The use of text support has been found to be key for aural word recognition [9,21], especially when it comes to the learning of a language with opaque orthography [22], and when learners are mostly exposed to written input.

The few experimental studies that have explored the development of L2 listening skills through captioned videos have indicated that the provision of text support improves learner-viewers' speech perception after relatively short interventions (1–2 episodes; 25–60 min). Still, it is uncertain whether the results from studies with university students may also be replicated with primary school students. In fact, the investigation by Tragant et al. [23] with fifth graders from Spain revealed that learners' exposure to 21 graded readers (i.e., texts adapted for foreign language learners) with audio support failed to enhance the development of L2 listening and reading skills, a finding that was attributed to the insufficient length of the treatment.

### *1.2. Captions and L2 Reading Skills Development*

L2 reading is a highly complex task that integrates lower- and higher-level reading processes (e.g., word recognition and global comprehension, respectively). The complexity of L2 reading might potentially explain why this activity has not proved very popular in foreign language contexts (e.g., [24,25]). This is a great limitation considering that practice in reading is crucial to show significant improvement in the development of this receptive language skill [16,26]. In L1 contexts, a handful of investigations with primary school learners have demonstrated that the use of L1 captions enhances the development of reading skills (e.g., [13–15]). Nonetheless, these findings from L1 contexts may not necessarily be translated to foreign language contexts. L1 and L2 readers differ in terms of language proficiency and their exposure to written texts, which are crucial factors for becoming familiar with L2 orthographic patterns and for automatizing lower-level reading skills [27]. In sum, the extent to which the use of onscreen text supports the development of L2 reading skills in foreign language settings is still unknown.

Up to now, the studies conducted with primary school learners in L1 contexts have demonstrated that the use of bimodal verbal input (i.e., audio and text) enhances the development of reading skills [13]. However, the processing patterns and the specific aspects that may benefit from this activity seem to depend on the extent to which lower-level reading skills are automatized. On the whole, the evidence has indicated that learners' exposure to captions improves word decoding skills [13–15]. However, at earlier stages of reading skills development (7–9 years old, 2nd grade), children seem to focus their attention on lower-level reading skills while processing captions, which is a factor that might hinder their capacity to focus on details or less relevant elements from the input [14]. As explained by Sadoski and Paivio [28], when the decoding process is effortful, learners devote greater attention to lower linguistic levels, leaving fewer cognitive resources available to process other elements, such as images and gestures, and to make referential connections. As Linebarger et al. [15] hypothesize, there seems to be a stage where the reading of captions is neither too challenging nor too easy to follow, so the use of text support may successfully aid comprehension and foster reading skills development.

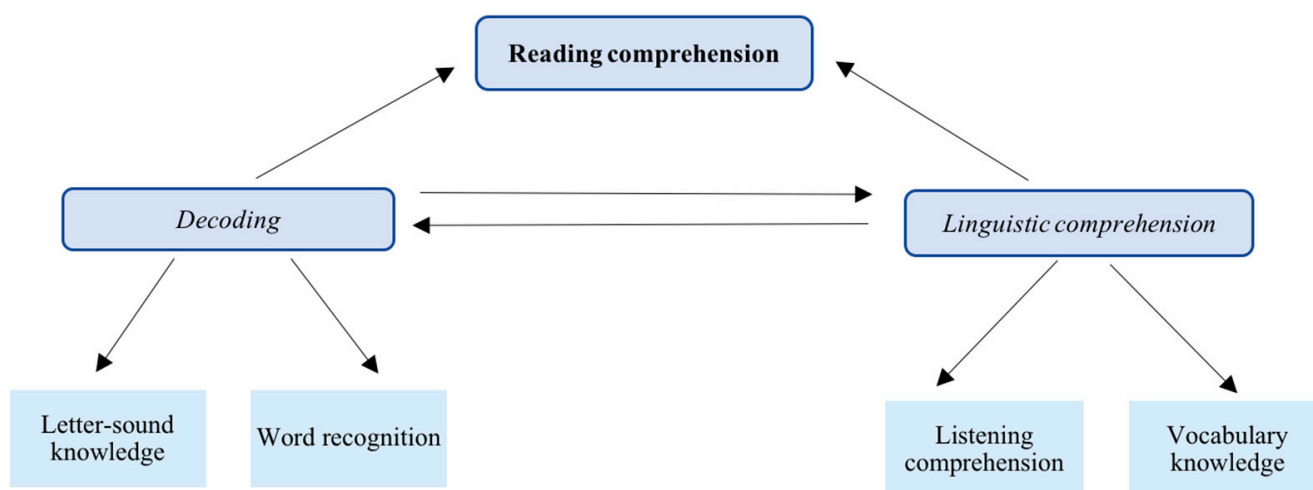
Furthermore, the reading aspects that may benefit from the use of onscreen text may also depend on the length of the treatment. The results obtained by Linebarger et al. [15] with second and third graders in the US (native and second-language learners) indicated that the students who were exposed to captions improved in terms of non-word reading (English patterns) but not in terms of oral reading fluency. In this regard, six episodes may not have been enough to lead to significant improvement. In other words, primary school learners might need intensive exposure to captions to significantly benefit from their viewing experience.

While it is true that the ultimate goal of reading instruction is to achieve high levels of comprehension, the literature suggests that the instruction and development of lower-level reading skills are crucial to attaining this objective [28]. Learners' development of both higher- and lower-level reading processes is key to building coherent mental representations [16,29,30]. With this in mind, it could also be stated that although captions support the development of lower-level reading skills at the expense of students' immediate viewing comprehension at early reading stages [14], the use of on-screen text might still be seen as a contribution to the learning process. That is to say, despite learners' greater efforts to cope with the speed of captions at early learning stages, the beneficial effects of captions on the automatization of orthographic and phonological processing may eventually result in higher levels of comprehension and motivation to read [22]. That being the case, captions may have the potential to break the vicious circle of low-achievers' reluctance to read [26] and, to a certain extent, counteract learners' lack of exposure to L2 print due to the complexity of L2 reading.

### *1.3. The Influence of Language-Related Factors*

#### *1.3.1. L1- vs. L2-Related Factors*

The extent to which L2 learners process audiovisual input with ease, and benefit from their viewing experience, seems to be strongly predicted by L2 proficiency level [31]. At lower proficiency levels, the processing of onscreen text seems to be more effortful, increasing the amount of time viewers spend on captions/subtitles [32,33]. This finding is to be expected considering the complexity of L2 reading and the fact that learners' ability to read texts with ease and a high level of comprehension appears to be mainly explained by L2-related factors [34,35]. Although the literature has consistently shown evidence of the influence of L1 reading skills on L2 reading [23,26,36,37], it has also been suggested that L2-related factors might be stronger predictors of L2 reading (e.g., [34,35,38,39]). The strong relationship between L2 reading and L2-related factors has been explained in terms of the Simple View of Reading model [40,41] (see Figure 1), which postulates that reading comprehension is mainly explained by word decoding and general oral language comprehension [35].



**Figure 1.** The modified Simple View of Reading.

### 1.3.2. L1 and L2 Reading Skills

To date, few researchers have focused their attention on the direct relationship between L2 receptive language skills (reading and listening) and L2 learning from captioned videos [42,43], and no one, to the best of our knowledge, has studied the specific influence of L1 reading skills in this regard. L1 and L2 reading skills may be particularly relevant in the case of young learners as their reading skills are still developing [17], also due to the effort required to cope with the speed of captions [32], as well as the need to integrate verbal and pictorial information to attain appropriate levels of comprehension [28,44]. Concerning the influence of L1 reading skills, the literature suggests that learners' L1 orthography may support and facilitate L2 reading to compensate for L2 knowledge gaps and lack of practice, as long as there is an overlap between the two systems [45]. Therefore, learners are thought to progressively assimilate and accommodate their linguistic infrastructure to the characteristics of the L2 [26,37,46], which is a process that relies on their L2 proficiency and familiarity with the characteristics of the target language [46]. Hence, one may expect that at least in the case of young L2 learners, both L1 and L2 reading skills might play a role in the processing of captions and in their ability to learn from them.

### 1.3.3. L2 Listening Skills

As for the role of L2 listening skills, the majority of studies on audiovisual input have focused on viewing comprehension or the development of listening skills rather than on the role of L2 listening in language learning from viewing, with few exceptions (e.g., [43,47]). Overall, it is widely accepted that L2 viewing without text support may be quite challenging for lower proficiency learners [48,49]. Therefore, one may assume that L2 proficiency, including L2 listening skills, may play a significant role in L2 learning from viewing without captions. Nonetheless, the scant evidence available indicates that listening skills may also predict the outcomes under the presence of captions, which is a factor that may not only be attributed to the input received through the aural channel but also to the general comprehension processes involved while viewing.

### 1.3.4. L2 Vocabulary Knowledge

The empirical evidence has demonstrated that a minimum level of vocabulary knowledge is required in order to show appropriate levels of comprehension in different modalities, namely aural and/or written (e.g., [50–52]). Once this threshold is surpassed, L2 learning is likely to occur due to the lower effort required in the processes of decoding and comprehension, which means that sufficient resources are available to notice unknown target language constructions [53,54]. While the lexical coverage that ensures appropriate levels of comprehension in viewing has been found to be less demanding than in reading-

only and listening-only conditions (80%, [50]) due to the presence of imagery, the majority of studies on audiovisual input have identified L2 vocabulary knowledge as a significant predictor of L2 learning from viewing in different age groups (e.g., [55,56]).

## 2. Materials and Methods

This investigation attempted to fill some of the gaps in the literature by determining whether the use of age-appropriate captioned videos was conducive to the development of L2 receptive skills in a group of primary school learners of English as a foreign language. In addition, this study investigated the role played by a series of language-related factors on the outcomes (see Table 1).

**Table 1.** Research design.

1. Pretest:		2. Treatment	3. Immediate Posttest:	4. Delayed Posttest:
Target Factors	Language-Related Factors		Target Factors	Target Factors
L2 reading efficacy	(L2 vocabulary knowledge, L1 reading efficacy, L2 listening skills)	11 episodes of the animated cartoon Charlie and Lola	L2 reading efficacy	L2 reading efficacy
L2 listening skills	(L2 vocabulary knowledge, L1 reading efficacy, L2 reading efficacy)		L2 listening skills	L2 listening skills

The research questions that guided this study were the following:

- (1) To what extent does the use of captioned videos foster the development of L2 listening skills and L2 reading efficacy in fourth and fifth graders?
- (2) To what extent do language-related factors influence the development of L2 listening skills and L2 reading efficacy? (i.e., L2 vocabulary knowledge, L2 listening skills, and L1 and L2 reading efficacy).

### 2.1. Participants

This study was conducted with a convenience sample of 96 L1-Spanish primary school learners in year 4 (aged 9–10;  $n = 47$ , 3 groups) and year 5 (aged 10–11;  $n = 49$ , 2 groups) from a private school in Chile. The outcomes of both year levels were compared due to the fast and marked changes experienced by children during middle childhood as regards their physical, socio-emotional, and cognitive development [57]. The number of female and male participants was comparable (girls = 47; boys = 49). The following requirements were considered for the inclusion of the participants in the statistical analyses: watching 100% of the episodes, doing at least a set of pre and posttests to assess their progress over time, and not being part of the group of students with special educational needs (who in fact received additional support to complete the activities). In addition, parents' consent was requested prior to the intervention, which was complemented by children's oral confirmation of their willingness to participate in the viewing experience.

Not all private and semi-private schools in Chile ensure high levels of proficiency through formal instruction, especially when schools are located outside the three metropolitan centers [58]. The participants from this school had received around six hours a week of formal L2 instruction since preschool. However, the amount of out-of-school contact with the target language reported in a questionnaire by the participants was either limited or non-existent, a factor that may explain their low level of proficiency (approximately pre-A1 according to the CEFR).

Two groups of students from the same school participated as control groups, which only completed a specific set of tests and attended their regular English lessons. Specifically, control group 1 (CG1;  $n = 16$  students, nine male and seven female, year 5) took the L2 listening tests (pretest and posttest; 5-week interval) a year before the intervention with the experimental groups (2020), while control group 2 (CG2;  $n = 17$  students, nine male and



eight female, year 4) took the L2 reading efficacy tests (pretest and posttest; 5-week interval) a year after the actual experiment (2022).

## 2.2. Treatment

The experimental groups watched 11 episodes of the animated cartoon Charlie and Lola [59], which was considered to be age- and content-appropriate. An important characteristic of this animated cartoon is that each episode lasts 10 min. Previous research suggested that after 10 min, young L2 learners may not be able to cope with the cognitive demands of the viewing task, in particular those who are less skilled readers [7]. Moreover, the characteristics of this animated cartoon seem to ensure appropriate levels of comprehension. Specifically, the dialogues contain sufficient visual support, and there is a clear connection between verbal input and the actions performed by the characters [60]. The linguistic analyses of the scripts (VocabProfiler on Lextutor; [61]) revealed that, on average, the episodes reached 91.3% vocabulary coverage at the K1 level. Based on the threshold proposed by previous research on vocabulary coverage with audiovisual input (80%; [50,51]), the analyses suggested that Charlie and Lola ensured appropriate levels of comprehension in low-proficiency learners. This was confirmed by a pilot group of fourth graders who reported that the episodes were suitable and motivating. Likewise, the results obtained in our pilot study are consistent with the outcomes obtained by Tragant and Pellicer-Sánchez with eye-tracking methodology [33]. By exploring fifth graders' eye movements while watching an episode of Charlie and Lola, the researchers concluded that despite the participants' higher level of attention to written input, they were fully capable of processing both text and images.

## 2.3. Instruments

Due to the pandemic, the instruments were carefully designed to be administered in pen-and-paper and online format (Google Forms). Although this investigation was conducted onsite with the experimental groups, the students who were put in quarantine had to complete some activities online through Microsoft Teams, which was the official platform used at the school. These online activities were organized either individually or in small groups (four students maximum). The participants were asked to keep their cameras on during the whole session.

### 2.3.1. Reading Efficacy in English and Spanish

The concept of reading efficacy (see [36]) integrates learners' reading speed (words read per minute = WPM) and comprehension, because both lower- and higher-level reading processes are thought to be equally important for discriminating high- and low-achievers [27]. Silent reading was selected over reading aloud because the former may better resemble the processing of captions and may be a more reliable instrument to test comprehension [29].

Several procedures were followed to ensure that the three texts selected to test reading efficacy in Spanish and English were comparable. As for Spanish reading efficacy (SR efficacy hereafter), the fiction texts (A, B, and C) were adapted from a set of sample materials [62] developed to train Chilean fourth graders for the national standardized test on reading skills. The ATOS readability formula, which measures text complexity as a function of average sentence length, average word length, and word difficulty level, indicated that the texts were appropriate for the sample groups. In addition, the six multiple-choice comprehension questions focused either on textually explicit/literal (four questions) or textually implicit information (two questions). Each test item consisted of a correct answer, three distractors, and the 'I don't know' option to prevent learners from guessing. Due to the participants' low L2 proficiency level, the non-fiction texts (A, B, and C) selected for the English reading efficacy tests (ER efficacy hereafter) were adapted from Pre-A1 starters' sample papers [63]. The assessment of text readability (i.e., Flesch–Kincaid reading ease and grade level) indicated that all the texts were easy to read. More specifically, we confirmed that the texts used at each testing time were comparable and suitable for the target groups.

The five multiple-choice comprehension questions followed the same structure as the SR efficacy tests. Four questions focused on textually explicit information while only one tested textually implicit information.

The measures obtained from these instruments were reading speed ( $\text{WPM} = [\text{n}^\circ \text{ of words in the text} / \text{number of seconds used to read the whole passage}] \times 60$ ) and comprehension. Each comprehension question was assigned one point. Then, the raw comprehension score was used to calculate the percentage of comprehension ( $\text{number of correct answers} \times 100 / \text{N}^\circ \text{ of questions}$ ). Finally, the formula used to calculate reading efficacy was  $([\text{WPM} \times \% \text{ comprehension}] / 100)$ . In view of the fact that reading speed (WPM) varied among the participants, there was no maximum score for this test.

Prior to their administration, these tests were pilot tested with 14 primary school learners from Chile and the conflictive items were improved before the intervention. With the experimental and control groups, the reading efficacy tests were administered in small sub-groups (four students maximum) to more accurately track reading speed (by using a stopwatch per child) and to ensure that the instructions were appropriately followed. First, the students were asked to read the texts at their own pace for comprehension purposes. This instruction was repeated several times to fulfill the main aim of this instrument. In addition, the instructions highlighted that the text had to be read only once. Having listened to the instructions, the students were explicitly told to start reading. For practical reasons, the texts contained red circles, which signaled the beginning and the end of the reading process. Specifically, the learners had to raise their hands when they reached the end. While answering the questions, the students did not have access to the text again.

### 2.3.2. L2 Listening Skills

Two sample Movers tests (paper A and paper B) [63] were implemented to measure learners' listening skills at the three testing times (i.e., listening for specific information). Paper A was previously pilot tested onsite with two groups of fifth graders from the same school (1.5 years before the actual experiment) to ensure its suitability. As a result, it was decided to distribute the test in two sessions.

The maximum score was 20 points. Correct answers were given one point, so each section was worth a total of five points. The test was administered onsite in pen-and-paper format. An online version was also developed in order to assess the control group, as well as the students that were in quarantine. Paper A was administered at pretest and delayed posttest, while paper B was given at posttest. The Cronbach alpha coefficient obtained for test A was 0.722 with the pilot group. As for the experimental groups, the Cronbach alpha values obtained for paper A were 0.655 at pretest, and 0.794 at delayed posttest, while the coefficient obtained for paper B at posttest was 0.682. Given that reliability analyses performed with a small number of items tend to lead to low values [64], and this test generated four main scores, the results suggest that these instruments' internal consistency is acceptable [65].

### 2.3.3. L2 Vocabulary Knowledge

The EFL picture vocabulary test assessed general vocabulary knowledge at the level of meaning recognition. This test was adapted from Puimège and Peters' [65] version of the Picture Vocabulary Size Test (PVST) created by Anthony and Nation [66] to assess young learners' vocabulary knowledge (L1 and L2 English speakers). The original instrument was pilot tested, and the results suggested that this test was not suitable for our context. Then, we designed an instrument with 50 items by keeping the same format. The target words, selected from the A2 key for schools' vocabulary list developed by Cambridge English Assessment [67], contained an equal number of items from the K1 and K2 frequency bands based on the analysis performed on Lextutor [61].

As for the testing procedures, the target words were uttered in isolation and then in a non-defining sentence [65]. These stimuli were simultaneously presented in written and oral form. The audio had been previously recorded by an English native speaker. Out

of four pictures, the students had to select the one that represented the meaning of the target word (A, B, C, or D). Additionally, the students had the possibility of selecting the 'I don't know' option to prevent guessing. The students had only 10 s to select the correct alternative for each testing item. The questions were presented through a video to ensure that the testing procedures were the same in all the experimental groups. In the online format, students could hear the audio and see the pictures on their form, while in the pen-and-paper format, a projector and speakers were used to show the video, while the students had to record their responses on an answer sheet.

The instrument was first pilot tested on six groups of EFL learners from Chile ( $N = 188$ ; Third-sixth graders). The Cronbach alpha coefficients obtained with the pilot groups were satisfactory (0.908 for K1 words, 0.898 for K2 words, and 0.898 for the whole test). In the experimental groups, the Cronbach alpha coefficients obtained were 0.866 for K1 words, 0.814 for K2 words, and 0.913 for the whole test. In addition, at the end of the pilot testing sessions, the EFL teachers answered a questionnaire (see [68]), which confirmed that the instrument was appropriate for the context. The pictures that were considered confusing were replaced by clearer options.

#### 2.4. Analyses

Data analyses were performed in SPSS v.25. First, we ran a series of ANOVAs and T-tests to ensure that the groups in each year level were comparable in terms of language-related factors. Then, Pearson's correlations were performed to study the relationships between variables (receptive L2 skills and language-related factors). The factors that were not normally distributed (Kolmogorov–Smirnov/Shapiro–Wilk =  $p < 0.05$ ) were square root (SQRT) transformed to run parametric tests. To measure learners' progress over time and the influence of language-related factors, a series of GLMMs were performed (generalized linear mixed models; L2 listening skills: binary logistic regressions, ER efficacy: linear models). These analyses included Satterthwaite approximation and robust covariances, which are suggested for small sample groups and unbalanced data. Additionally, multiple linear regressions were run to assess the influence of language-related factors and to calculate the specific contribution of each variable on ER efficacy. Prior to the calculation of GLMMs and multiple linear regressions, we assessed collinearity between variables (tolerance  $> 0.3$ ; VIF  $< 3.33$ ). As regards GLMMs, model fit was estimated through AIC (Akaike information criterion). The GLMMs consisted of a compound-symmetry structure with student identification as subjects and time as a repeated measure.

### 3. Results

#### 3.1. Preliminary Analyses

Independent-sample T-tests were calculated to compare fourth and fifth graders in language-related factors. The results demonstrated that at pretest, fifth graders scored significantly higher than fourth graders in vocabulary knowledge ( $t(86) = 3.195$ ,  $p = 0.006$ ,  $r = 0.32$ ), L2 listening skills ( $t(94) = 17.921$ ,  $p < 0.001$ ,  $r = 0.87$ ), and ER efficacy ( $t(107) = 4.320$ ,  $p < 0.001$ ,  $r = 0.38$ ). However, the difference between the two year levels in SR efficacy only approached significance ( $t(106) = 1.923$ ,  $p = 0.057$ ,  $r = 0.18$ ) (see Tables 2 and 3).

**Table 2.** L2 vocabulary knowledge and SR efficacy: descriptive statistics.

	Year Level			
	Year 4		Year 5	
	M	(SD)	M	(SD)
Vocabulary knowledge	14.47	(6.28)	20.5	(10.81)
SR efficacy	79.15	(40.17)	94.85	(46.33)



**Table 3.** L2 listening skills and ER efficacy: descriptive statistics.

		Listening Pretest		Listening Posttest		Listening Delayed	
		Mean	(SD)	Mean	(SD)	Mean	(SD)
Group	Year 4	8.46	(3.35)	11.19	(3.46)	11.15	(3.78)
	Year 5	11.31	(4.51)	13.63	(3.81)	14.16	(3.85)
	CG1-fifth	11.56	(3.12)	10.06	(3.09)		
		ER Efficacy Pretest		ER Efficacy Posttest		ER Efficacy Delayed	
		Mean	(SD)	Mean	(SD)	Mean	(SD)
Group	Year 4	45.00	(30.36)	68.98	(46.17)	80.81	(33.03)
	Year 5	72.72	(34.82)	109.59	(52.17)	124.29	(48.46)
	CG2-fourth	50.06	(31.24)	42.53	(21.18)		

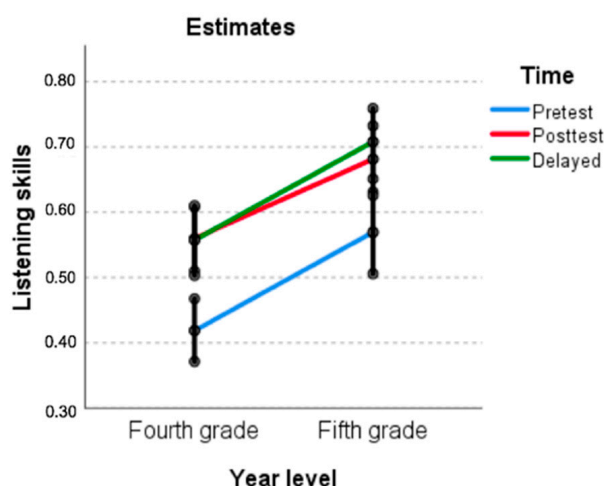
### 3.2. Development of L2 Listening Skills from Captioned-Video Viewing and the Influence of Language-Related Factors

We ran a compound symmetry structure GLMM (binary logistic regression) with student identification as subjects and time as a repeated measure to assess each year level's progress over time. The model was built with learners' scores at the three testing times by setting 20 (maximum score) as the denominator. The fixed effects included in the analyses were time, year level, and their interaction. The results yielded significant main effects for time ( $F(2,190) = 63.966, p < 0.001$ ) and year level ( $F(1,89) = 15.844, p < 0.001$ ), but a non-significant interaction between these two factors. As shown in Table 4, the experimental groups showed significant progress from pretest to posttest, and from pretest to delayed posttest, regardless of their year level. The results also showed that learners' scores did not significantly decrease from posttest to delayed posttest. The significant effects of year level were associated with the fact that the higher performance of fifth graders was kept over time. On the whole, the results indicated that the treatment was similarly beneficial for both year levels (see Figure 2).

**Table 4.** Listening skills: Time pairwise contrasts.

Time Pairwise Contrasts	Contrast Estimate	Std. Error	t	df	Adj. Sig.	95% CI	
						Lower	Upper
Pretest–posttest	−0.129	0.014	−9.224	241	0.000	−0.162	−0.095
Pretest–delayed	−0.142	0.015	−9.785	197	0.000	−0.175	−0.109
Posttest–delayed	−0.014	0.015	−0.894	148	0.373	−0.044	0.016

The sequential Bonferroni-adjusted significance level is 0.05. Confidence interval bounds are approximate.

**Figure 2.** Listening skills: the trajectory of each year level over time.

To compare the experimental groups' performance with the control group (CG1-fifth) at pretest and posttest, we ran a compound symmetry structure GLMM (binary logistic regression). The results revealed significant effects for group ( $F(2,135) = 6.768, p = 0.002$ ), time ( $F(1,97) = 16.887, p < 0.001$ ), and their interaction ( $F(2,102) = 14.495, p < 0.001$ ). Specifically, the Bonferroni-adjusted results indicated that the groups significantly improved from pretest to posttest ( $p < 0.001$ ), except for the control group, which in fact scored lower at posttest (see Table 5).

**Table 5.** Listening skills: Time pairwise contrasts by class (pretest and posttest).

Groups	Time Pairwise Contrasts	Contrast Estimate	Std. Error	t	df	Adj. Sig.	95% CI	
							Lower	Upper
Fourth grade	Pretest–posttest	−0.141	0.021	−6.853	95	0.000	−0.182	−0.100
Fifth grade	Pretest–posttest	−0.115	0.019	−6.155	143	0.000	−0.152	−0.078
CG1-fifth	Pretest–posttest	0.075	0.036	2.105	89	0.038	0.004	0.146

The sequential Bonferroni-adjusted significance level is 0.05. Confidence interval bounds are approximate.

As for the influence of language-related factors on the development of L2 listening skills, a series of generalized linear mixed models (binary logistic regression) with repeated measures (time) compound-symmetry structure were calculated. To this end, the following factors were entered into the model: L2 vocabulary knowledge, SR efficacy, ER efficacy, year level, time, and all possible two-way interactions. The best-fitted model was determined by a backward elimination procedure. The results yielded significant effects for vocabulary knowledge ( $F(1,82) = 56.549, p < 0.001$ ), SR efficacy ( $F(1,119) = 6.754, p = 0.011$ ), and time ( $F(2,173) = 59.388, p < 0.001$ ). The exponential coefficients in Table 6 indicate that the odds of obtaining a correct response in the listening test increased by 69% per each additional point on the EFL vocabulary test, and by 6.8% per each additional point in SR efficacy.

**Table 6.** Listening skills: best fitted model obtained to assess the influence of language-related factors on learners' scores.

Model Term	Coef	Std. Error	t	Sig.	95% CI		Exp (Coef)	95% CI for Exp(Coef)	
					Lower	Upper		Lower	Upper
Intercept	−2.166	0.2748	−7.882	0.000	−2.711	−1.621	0.115	0.066	0.198
Vocabulary knowledge	0.528	0.0702	7.520	0.000	0.388	0.667	1.695	1.474	1.949
SR efficacy	0.066	0.0252	2.599	0.011	0.016	0.115	1.068	1.016	1.122
Pretest	−0.604	0.0669	−9.029	0.000	−0.736	−0.472	0.547	0.479	0.624
Posttest	−0.018	0.0730	−0.252	0.802	−0.163	0.126	0.982	0.850	1.134
Delayed posttest	0 <sup>b</sup>								

Probability distribution: binomial. Link function: logit. <sup>b</sup> This coefficient is set to zero because it is redundant.

### 3.3. Development of ER Efficacy from Captioned-Video Viewing and the Influence of Language-Related Factors

To compare the performance of fourth and fifth graders over time, we calculated a compound symmetry structure GLMM (linear model) with student identification as subjects, and time as a repeated measure. The model was fitted with learners' scores at the three testing times and the following fixed factors: time, year level, and their interaction. The results yielded significant main effects for year level ( $F(1,94) = 27.711, p < 0.001$ ) and time ( $F(2,166) = 72.697, p < 0.001$ ), but not for their interaction. However, the interaction was kept in the model to further observe each year level's outcomes (see Figure 3 and Table 7).

On the whole, the experimental groups showed significant gains from pretest to posttest, and from pretest to delayed posttest, regardless of their year level. Moreover, the results indicated that learners' scores significantly increased from posttest to delayed posttest. The significant effects of year level confirmed that fifth graders consistently outperformed fourth graders over time. As shown in Table 7, fifth graders seemed to obtain slightly higher gains from the treatment.

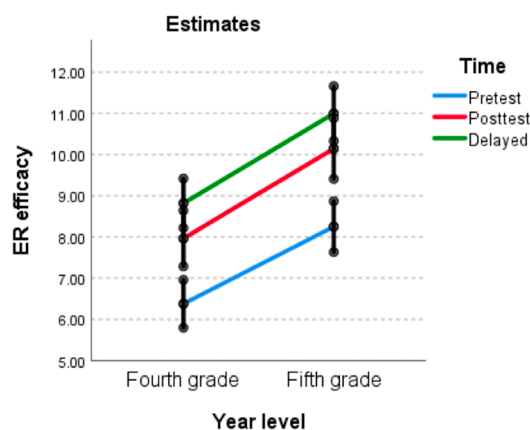


Figure 3. ER efficacy: the trajectory of each year level over time.

Table 7. ER efficacy: time pairwise contrasts by year level.

Year Level	Time Pairwise Contrasts	Contrast Estimate	Std. Error	t	df	Adj. Sig.	95% CI	
							Lower	Upper
Fourth grade	Pretest–posttest	−1.584	0.256	−6.199	261	0.000	−2.160	−1.008
	Pretest–delayed	−2.440	0.300	−8.145	185	0.000	−3.164	−1.716
	Posttest–delayed	−0.856	0.302	−2.831	178	0.005	−1.453	−0.259
Fifth grade	Pretest–posttest	−1.893	0.275	−6.884	201	0.000	−2.514	−1.272
	Pretest–delayed	−2.744	0.349	−7.855	96	0.000	−3.596	−1.893
	Posttest–delayed	−0.851	0.304	−2.804	146	0.006	−1.451	−0.251

The sequential Bonferroni-adjusted significance level is 0.05. Confidence interval bounds are approximate.

To compare the performance of the two year levels and the control group, a new compound symmetry structure GLMM (linear model) was calculated with student identification as subjects, and time as a repeated measure. To this end, we fitted learners’ listening scores as the target variable, and entered group, time, and their interaction into the model as fixed factors. The results yielded significant effects for time ( $F(1,120) = 29.693, p < 0.001$ ), group ( $F(2,130) = 16.853, p < 0.001$ ), and a significant interaction between time and group ( $F(2,116) = 11.889, p < 0.001$ ). Specifically, the Bonferroni-adjusted results revealed that both year levels significantly improved from pretest to posttest ( $p < 0.001$ ), while the control group obtained similar scores at the two testing times (see Table 8).

Table 8. ER efficacy: time pairwise contrasts by group (pretest and posttest).

Groups	Time Pairwise Contrasts	Contrast Estimate	Std. Error	t	df	Adj. Sig.	95% CI	
							Lower	Upper
Fourth grade	Pretest–posttest	−1.584	0.256	−6.199	122	0.000	−2.090	−1.078
Fifth grade	Pretest–posttest	−1.892	0.275	−6.877	95	0.000	−2.438	−1.346
CG2-fourth	Pretest–posttest	0.443	0.411	1.079	134	0.282	−0.369	1.256

The sequential Bonferroni-adjusted significance level is 0.05. Confidence interval bounds are approximate.

Additionally, a new model was built to assess each year level’s progress in silent reading speed (number of words read per minute). Specifically, we ran a series of repeated measures (time) compound-symmetry structure GLMMs (linear model) with learners’ scores in silent reading speed as the target factor, and the following variables as fixed factors: year level, time, and their interaction. The results revealed significant effects for year level ( $F(1,89) = 21.346, p < 0.001$ ) and time ( $F(2,175) = 20.950, p < 0.001$ ), while their interaction did not reach significance ( $F(2,175) = 0.914, p = 0.403$ ). Nonetheless, this interaction was kept in the model to further explore the progress of each year level. As shown in Table 9, both year levels significantly improved from pretest to posttest, and from pretest to delayed posttest ( $p < 0.05$ ). However, the results suggest that fifth graders

showed greater progress in reading speed over time, especially from pretest to delayed posttest, and from posttest to delayed posttest.

**Table 9.** ER efficacy: pairwise contrasts of the outcomes obtained by each group (WPM) over time.

Year Level	Pairwise Contrasts	Contrast Estimate	Std. Error	t	df	Adj. Sig.	95% CI	
							Lower	Upper
Fourth grade	Pretest–posttest	−0.561	0.217	−2.588	138	0.021	−1.053	−0.070
	Pretest–delayed	−0.718	0.211	−3.398	191	0.002	−1.229	−0.208
	Posttest–delayed	−0.157	0.223	−0.702	149	0.484	−0.598	0.284
Fifth grade	Pretest–posttest	−0.743	0.192	−3.872	237	0.000	−1.175	−0.310
	Pretest–delayed	−1.119	0.208	−5.376	188	0.000	−1.622	−0.616
	Posttest–delayed	−0.377	0.209	−1.804	187	0.073	−0.788	0.035

The sequential Bonferroni-adjusted significance level is 0.05. Confidence interval bounds are approximate.

As regards the influence of language-related factors, we calculated a series of repeated-measures (time) compound-symmetry structure GLMMs (linear models). To this aim, we entered ER efficacy (students' scores at the three testing times) as the target factor, and the following variables as fixed effects: year level, time, L2 listening skills, SR efficacy, L2-vocabulary knowledge, and all possible two-way interactions. By following a step-back procedure, the non-significant interactions and main effects were removed from the model one by one until the best-fitted model was obtained (see Table 10). The results revealed significant main effects for year level ( $F(1,53) = 9.186, p = 0.004$ ), listening skills ( $F(1,66) = 4.525, p = 0.037$ ), L2 vocabulary knowledge ( $F(1,80) = 29.818, p < 0.001$ ), SR efficacy ( $F(1,104) = 11.395, p = 0.001$ ), and time ( $F(2,121) = 4.302, p = 0.016$ ). In addition, the results yielded a significant interaction between L2 listening skills and time ( $F(2,154) = 4.351, p = 0.015$ ).

**Table 10.** ER efficacy: The influence of L2-related factors on learners' scores.

Model Term	Coef	Std. Error	t	Sig.	95% CI	
					Lower	Upper
Intercept	3.439	0.8089	4.251	0.000	1.828	5.049
Year level	−0.934	0.3082	−3.031	0.004	−1.552	−0.316
Listening skills	0.116	0.0561	2.068	0.041	0.005	0.227
Vocabulary knowledge	0.970	0.1777	5.461	0.000	0.617	1.324
SR efficacy	0.212	0.0628	3.376	0.001	0.087	0.337
Pretest	−1.694	0.6257	−2.707	0.008	−2.936	−0.452
Posttest	−1.464	0.5762	−2.541	0.012	−2.605	−0.323
Delayed	0 <sup>b</sup>					
Listening skills *[pretest]	−0.093	0.0538	−1.728	0.086	−0.199	0.013
Listening skills *[posttest]	0.055	0.0487	1.125	0.262	−0.041	0.151
Listening skills *[delayed]	0 <sup>b</sup>					

Probability distribution: normal. Link function: identity. <sup>b</sup> This coefficient is set to zero because it is redundant.  
\* Interaction.

To calculate the contribution of each factor on students' ER efficacy scores at pretest, posttest and delayed posttest, we ran multiple linear regressions for each testing time. The predictor variables included were as follows: listening skills, vocabulary knowledge, and SR efficacy. At pretest, L2 listening skills were not found to contribute significantly ( $p > 0.05$ ), so this factor was removed from the analysis. The results indicated that vocabulary knowledge and SR efficacy predicted 44% of the variance ( $F(2,85) = 36.404, p < 0.001, R^2 = 0.449$ ). Specifically, the standard coefficients indicated that vocabulary knowledge was the strongest predictor ( $\beta = 47\%, p < 0.001$ ), followed by SR efficacy ( $\beta = 31\%, p = 0.001$ ). At posttest, the results indicated that listening skills, vocabulary knowledge, and SR efficacy predicted 57% of the variance ( $F(3,83) = 39.437, p < 0.001, R^2 = 0.573$ ). The standard coefficients indicated that vocabulary knowledge was the strongest predictor ( $\beta = 44\%, p < 0.001$ ),

followed by listening skills ( $\beta = 27\%$ ,  $p = 0.007$ ) and SR efficacy ( $\beta = 17\%$ ,  $p = 0.042$ ). Finally, at delayed posttest, the results revealed that SR efficacy was no longer significant ( $p = 0.100$ ), so this factor was removed from the analysis. The results showed that vocabulary knowledge and listening skills predicted 48% of the variance ( $F(2,76) = 38.210$ ,  $p < 0.001$ ,  $R^2 = 0.488$ ). Again, the standard coefficients indicated that vocabulary knowledge was a stronger predictor ( $\beta = 45\%$ ,  $p < 0.001$ ) than L2 listening skills ( $\beta = 32\%$ ,  $p = 0.004$ ).

#### 4. Discussion

This study aimed to determine the extent to which primary school learners may benefit from captioned video viewing with regard to the development of L2 listening skills and ER efficacy. In addition, we studied the influence of language-related factors on the outcomes. With respect to the first research question, the results indicated that in comparison with the control groups, the experimental groups obtained significant gains from the treatment in terms of L2 listening skills and ER efficacy. As for L2 listening skills, the results seemed to confirm that the use of captions enhanced bottom-up processing after a relatively short intervention (11 episodes), as previous studies have also found [9–11]. Likewise, the results concerning the development of ER efficacy suggested that the findings obtained by studies conducted in L1 settings (e.g., [13–15]) may be extrapolated to foreign language contexts. Additionally, the comparisons between year levels regarding their progress in silent reading speed indicated that both year levels significantly improved over time, although fifth graders were found to obtain greater gains in this regard. This outcome seems to indicate that learners' progress in ER efficacy may not only be associated with their improvement in silent reading speed but also with their capacity to devote fewer attentional resources to text decoding to improve their levels of comprehension [28,30]. As the literature suggests, the automatization of lower-level reading skills requires plenty of exposure to print [16]. Thus, fourth graders might need a higher number of episodes to show greater gains in silent reading speed [15].

The fact that the use of captioned videos enhanced the development of both receptive language skills may not be surprising, given that reading and listening have been found to have a bidirectional relationship [26,35,38,41], which may have been enhanced by learners' simultaneous exposure to both modalities [38]. As for the second research question, the results indicated that learners' outcomes were influenced by both L1 and L2-related factors. Specifically, their progress in L2 listening skills was found to be significantly predicted by L2 vocabulary knowledge and SR efficacy, while their performance in ER efficacy over time was significantly explained by vocabulary knowledge, L2 listening skills, and SR efficacy. However, when assessing the exact contribution of each variable, the findings suggested that learners' performance in L2 listening and reading skills over time was mainly explained by L2-related factors, particularly L2 vocabulary knowledge [34,35]. The finding that indicated that reading efficacy in Spanish but not in English influenced the development of L2 listening skills may be explained by the target participants' limited practice in L2 reading and viewing. Thus, in line with the literature, the outcomes suggest that the participants relied on their L1 linguistic infrastructure to process onscreen text [26,46]. In other words, young and low-proficiency learners may rely on L1 reading skills to compensate, to a certain extent, for their L2 knowledge gaps and lack of practice [39,45].

In general, the treatment appeared to be similarly beneficial for both year levels. Although it is true that fifth graders consistently outperformed fourth graders at the three testing times in L2 listening skills, the results may be attributed to their significantly higher proficiency level, and possibly to their higher cognitive development [12,17]. A similar picture was observed when assessing learners' performance in ER efficacy over time. However, as mentioned above, fifth graders appeared to show marginally higher improvement in ER efficacy, especially as regards silent reading speed. This outcome may be associated with their stronger L1 literacy skills [17], which may have allowed them to cope with the speed of captions and encouraged them to stay on the reading task



while viewing. Fifth graders may have been better equipped to rely on their L1 linguistic infrastructure to cope with the L2 input demands [26,37].

As for learners' trajectory from posttest to delayed posttest, the results indicated that their scores in L2 listening skills did not show significant variability. Yet, fifth graders showed a slight improvement between these two testing times. Similarly, fifth graders' scores in ER efficacy presented a marginally significant increase. While these outcomes may well be attributed to test effects, it may also be the case that the intervention encouraged fifth graders to watch videos at home. Prior research with adolescents has demonstrated that this type of intervention may have a positive impact on learners' viewing habits (e.g., [69]). In fact, the data collected in the present investigation by means of an interview with sample participants suggested that fifth graders were more open to experimenting with viewing and the use of onscreen text as a result of the intervention (see [68]). Nonetheless, more evidence needs to be gathered to confirm these outcomes. Future studies should explore further the extent to which young learners' viewing experiences at school may have an effect on their viewing habits at home [8].

The findings of the present study do not coincide with those of the study conducted by Tragant et al. [23] with fifth graders. In that study, learners' exposure to 21 graded readers (with and without audio support) was not conducive to significant gains in listening or reading skills in comparison with a control group. As Tragant et al. [23] explained, the length of the intervention may have been insufficient to observe significant gains from learners' exposure to graded readers, because the development of receptive language skills requires a great deal of practice [16], particularly in the case of young L2 learners, who have been found to be less efficient [18]. A key difference between the two studies was that our participants' viewing experience was limited or practically non-existent, while the students in Tragant et al.'s [23] investigation were already familiar with graded readers in L1 and L2. Thus, the gains in our study might be the result of the sudden increase in the participants' exposure to captioned videos (11 viewing sessions). Furthermore, the conflicting results may also be associated with the presence of moving images and the fact that the participants in the present study were not able to control the viewing process as had been the case in Tragant et al.'s [23] investigation, where each child could manipulate the audio (e.g., pause) and read the books at their own pace. Thus, learners' gains in L2 listening skills and ER efficacy may have been enhanced by the supporting role of imagery in terms of comprehension and the greater effort involved in the processing of captions during screen exposure. Further work is required to confirm whether the implementation of relatively short interventions that include learners' simultaneous exposure to audio and text may foster the development of receptive L2 skills. In summary, the comparison between reading-while-listening and captioned videos might also be a fruitful area for further research. By the same token, the use of eye-tracking methodologies would be of great help to explore young learners' processing patterns while reading static and dynamic texts with audio support, as well as to compare different age or proficiency groups [33].

Perhaps the main limitation of our investigation was that our data were collected during the COVID-19 pandemic, because our decisions were restricted by the contingency measures taken by the Chilean government. In 2021, few schools implemented in-person classes, and most of them were reluctant to welcome researchers into their classrooms. Therefore, we were unable to implement this research design at a different school to increase our sample size and/or compare learners' outcomes as a function of the characteristics of their language program. In addition, the present study measured gains after a relatively short intervention, because we worked under the threat of going online at any moment. Hence, we prioritized the completion of the experiment over the measurement of learners' outcomes after a higher number of sessions.

## 5. Conclusions and Pedagogical Implications

Taken together, the evidence obtained in this investigation supports the use of captioned videos with primary-school learners to foster the development of receptive L2 skills.

The results seem to demonstrate that fourth graders' lower proficiency level (in L1 and L2) and lack of exposure to the target language did not prevent them from benefiting from the treatment. Thus, the results lead to the conclusion that the use of captioned videos in the L2 classroom may help learners cope with the challenges entailed in the implementation of listening-only and reading-only activities [19,20] and may break down the potential vicious circle of low achievers' reluctance to complete reading and listening activities [26]. As stated above, young learners' exposure to the target language is a fundamental aspect of their language learning process [17,18]; therefore, the systematic use of captioned videos in the L2 classroom might enhance the outcomes of foreign language learning programs implemented in input-limited contexts.

**Author Contributions:** Conceptualization, D.A. and C.M.; methodology, D.A. and C.M.; analyses, D.A.; investigation, D.A. and C.M.; resources, D.A. and C.M.; data curation, D.A.; writing—original draft preparation, D.A.; writing—review and editing, D.A. and C.M.; supervision, C.M.; project administration, C.M. All authors have read and agreed to the published version of the manuscript.

**Funding:** This research was funded by the APIF scholarship granted by the University of Barcelona, and the grant PID2019-110594GB-I00 from the Spanish Ministry of Science, Innovation and Universities.

**Institutional Review Board Statement:** The study was conducted in accordance with the ethical principles proposed by the Research Ethics Committee at the University of Barcelona (2020-22). The study protocol guaranteed good practices in data collection, anonymization, processing, and storage. The study respected the privacy and confidentiality of all the participants involved. Apart from ensuring that parents' informed consent was provided, we requested students' verbal confirmation and respected their actual willingness to complete the activities implemented at school.

**Informed Consent Statement:** Informed consent was obtained from all subjects involved in the study.

**Data Availability Statement:** The data that support the findings of this study are available from the corresponding author, upon reasonable request.

**Conflicts of Interest:** The authors declare no conflict of interest.

## References

1. Montero Perez, M. Second or foreign language learning through watching audio-visual input and the role of on-screen text. *Lang. Teach.* **2022**, *55*, 163–192. [\[CrossRef\]](#)
2. Muñoz, C. Audiovisual input in L2 learning. *Lang. Interact. Acquis.* **2022**, *13*, 125–143. [\[CrossRef\]](#)
3. Montero Perez, M.; Rodgers, M. Video and language learning. *Lang. Learn. J.* **2019**, *47*, 403–406. [\[CrossRef\]](#)
4. Black, S. Subtitles as a tool to boost language learning? Children's views and experiences of watching films and television programmes in other languages with interlingual subtitles. *J. Audiovis. Transl.* **2022**, *5*, 73–93.
5. Lindgren, E.; Muñoz, C. The influence of exposure, parents, and linguistic distance on young European learners' foreign language comprehension. *Int. J. Multiling.* **2013**, *10*, 105–129. [\[CrossRef\]](#)
6. Muñoz, C.; Cadierno, T.; Casas, I. Different starting points for English language learning: A Comparative Study of Danish and Spanish Young Learners. *Lang. Learn.* **2018**, *68*, 1076–1109. [\[CrossRef\]](#)
7. Marzá, A.; Torralba, G. Incidental language learning through subtitled cartoons: Is it possible in a dubbing country? In *Subtitles and Language Learning. Principles, Strategies and Practical Experiences*; Gambier, Y., Caimi, A., Mariotti, C., Eds.; Peter Lang: Bern, Switzerland, 2015; pp. 199–219.
8. Webb, S. Extensive viewing: Language learning through watching television. In *Language Learning beyond the Classroom*; Nunan, D., Richards, J., Eds.; Routledge: New York, NY, USA, 2015; pp. 159–168.
9. Birulés-Muntané, J.; Soto-Faraco, S. Watching subtitled films can help learning foreign languages. *PLoS ONE* **2016**, *11*, e0158409. [\[CrossRef\]](#)
10. Charles, T.; Trenkic, D. Speech segmentation in a second language: The role of bi-modal input. In *Subtitles and Language Learning. Principles, Strategies and Practical Experiences*; Gambier, Y., Caimi, A., Mariotti, C., Eds.; Peter Lang: Bern, Switzerland, 2015; pp. 173–197.
11. Mitterer, H.; McQueen, J.M. Foreign subtitles help but native-language subtitles harm foreign speech perception. *PLoS ONE* **2009**, *4*, e7785. [\[CrossRef\]](#)
12. Muñoz, C. Symmetries and asymmetries of age effects in naturalistic and instructed L2 learning. *Appl. Linguist.* **2008**, *24*, 578–596. [\[CrossRef\]](#)
13. Kothari, B.; Takeda, J.; Joshi, A.; Pandey, A. Same language subtitling: A butterfly for literacy? *Int. J. Lifelong Educ.* **2002**, *21*, 55–66. [\[CrossRef\]](#)

14. Linebarger, D. Learning to read from television: The effects of using captions and narration. *J. Educ. Psychol.* **2001**, *93*, 288–298. [[CrossRef](#)]
15. Linebarger, D.; Taylor Piotrowski, J.; Greenwood, C. On-screen print: The role of captions as a supplemental literacy tool. *J. Res. Read.* **2010**, *33*, 148–167. [[CrossRef](#)]
16. Grabe, W.; Stoller, F. *Teaching and Researching Reading*, 3rd ed.; Routledge: New York, NY, USA, 2020.
17. Holmes, B.; Myles, F. White Paper: Primary Languages Policy in England. The Way Forward. RiPL. 2019. Available online: [www.ripl.uk/policy/](http://www.ripl.uk/policy/) (accessed on 15 May 2022).
18. Muñoz, C. *Age and the Rate of Foreign Language Learning*; Multilingual Matters: Clevedon, UK, 2006.
19. Kormos, J.; Kosak, M.; Pizorn, K. The role of low-level first language skills in second language reading, reading-while-listening and listening performance: A study of young dyslexic and non-dyslexic language learners. *Appl. Linguist.* **2019**, *40*, 834–858. [[CrossRef](#)]
20. Newton, J.; Nation, I.S.P. Extensive listening. In *Teaching ESL/EFL Listening and Speaking*, 2nd ed.; Newton, J., Nation, I.S.P., Eds.; Routledge: New York, NY, USA, 2021.
21. Bird, S.A.; Williams, J.N. The effect of bimodal input on implicit and explicit memory: An investigation into the benefits of within-language subtitling. *Appl. Psycholinguist.* **2002**, *23*, 509–533. [[CrossRef](#)]
22. Toscano-Fuentes, C.M.; Julián de Vega, C. Videos musicales en el aula de inglés de primaria para la mejora de la fluidez lectora. *Tejuelo* **2018**, *28*, 43–66.
23. Tragant, E.; Llanes, À.; Pinyana, À. Linguistic and non-linguistic outcomes of a reading-while-listening program for young learners of English. *Read. Writ.* **2019**, *32*, 819–838. [[CrossRef](#)]
24. Muñoz, C. Boys like games and girls like movies. Age and gender differences in out-of-school contact with English. *Rev. Española De Lingüística Apl.* **2020**, *33*, 171–201.
25. Peters, E. The effect of out-of-class exposure to English language media on learners' vocabulary knowledge. *Int. J. Appl. Linguist.* **2018**, *169*, 142–168.
26. Birch, B.; Fulop, S. *English L2 Reading: Getting to the Bottom*, 4th ed.; Routledge: New York, NY, USA, 2021.
27. Grabe, W. *Reading in a Second Language. Moving from Theory to Practice*; Cambridge University Press: Cambridge, UK, 2009.
28. Sadoski, M.; Paivio, A. A dual theoretical model of reading. In *Theoretical Models and Processes of Reading*; Alvermann, D.E., Unrau, N.J., Ruddell, R.B., Eds.; International Reading Association: Newark, NJ, USA, 2013; pp. 886–922.
29. Alderson, J.C.; Haapakangas, E.-L.; Huhta, A.; Nieminen, L.; Ullakonoja, R. *The Diagnosis of Reading in a Second or Foreign Language*; Routledge: New York, NY, USA, 2015.
30. Nassaji, H. The role and importance of lower-level processes in second language reading. *Lang. Teach.* **2015**, *47*, 1–37. [[CrossRef](#)]
31. Gesa, F.; Miralpeix, I. Effects of watching subtitled TV series on foreign language learning. Does proficiency level matter? In *Foreign Language Learning in the Digital Age. Theory and Pedagogy for Developing Literacies*; Lütge, C., Ed.; Routledge: Oxford, UK, 2022; pp. 159–173.
32. Muñoz, C. The role of age and proficiency in subtitle reading: An eye tracking study. *System* **2017**, *67*, 77–86. [[CrossRef](#)]
33. Tragant, E.; Pellicer-Sánchez, A. Young EFL learners' processing of multimodal input: Examining learners' eye movements. *System* **2019**, *80*, 212–223. [[CrossRef](#)]
34. Alderson, J.C.; Huhta, A.; Nieminen, L. Characteristics of weak and strong readers in a foreign language. *Mod. Lang. J.* **2016**, *100*, 853–879. [[CrossRef](#)]
35. Sparks, R.L. Identification and characteristics of strong, average, and weak foreign language readers: The simple view of reading model. *Mod. Lang. J.* **2021**, *105*, 507–525. [[CrossRef](#)]
36. Llanes, A. Reading in English as a foreign language: Examining differences in reading speed, comprehension, efficacy and L1 cross-linguistic influence across grades. *Investig. Sobre Lect.* **2018**, *9*, 1–49.
37. Perfetti, C.A.; Liu, Y.; Fiez, J.; Nelson, J.; Bolger, D.J.; Tan, L. Reading in two writing systems: Accommodation and assimilation of the brain's reading network. *Biling. Lang. Cogn.* **2007**, *10*, 131–146. [[CrossRef](#)]
38. Verhoeven, L.; van Leeuwe, J. The simple view of second language reading throughout the primary grades. *Read. Writ.* **2012**, *25*, 1805–1818. [[CrossRef](#)]
39. Yamashita, J. Mutual compensation between L1 reading ability and L2 language proficiency in L2 reading comprehension. *J. Res. Read.* **2002**, *25*, 81–95. [[CrossRef](#)]
40. Gough, P.; Tunmer, W. Decoding, reading, and reading disability. *Remedial Spec. Educ.* **1986**, *7*, 6–10. [[CrossRef](#)]
41. Tunmer, W.E.; Chapman, J.W. The simple view of reading redux. *J. Learn. Disabil.* **2012**, *45*, 453–466. [[CrossRef](#)]
42. Muñoz, C.; Pattemore, A.; Avello, D. Exploring repeated captioning viewing as a way to promote vocabulary learning: Time lag between repetitions and learner factors. *Comput. Assist. Lang. Learn.* **2022**. [[CrossRef](#)]
43. Suárez, M.; Gesa, F. Learning vocabulary with the support of sustained exposure to captioned video: Do proficiency and aptitude make a difference? *Lang. Learn. J.* **2019**, *47*, 479–496. [[CrossRef](#)]
44. Pellicer-Sánchez, A.; Tragant, E.; Conklin, K.; Rodgers, M.; Serrano, R.; Llanes, A. Young learners' processing of multimodal input and its impact on reading comprehension. An eye tracking study. *Stud. Second Lang. Acquis.* **2020**, *42*, 577–598. [[CrossRef](#)]
45. Birch, B. *English L2 Reading: Getting to the Bottom*, 3rd ed.; Routledge: New York, NY, USA, 2015.
46. Jiang, J.; Ouyang, J.; Liu, H. Interlanguage: A perspective of quantitative linguistic typology. *Lang. Sci.* **2019**, *74*, 85–97. [[CrossRef](#)]

47. Pattemore, A.; Muñoz, C. Learning L2 constructions from captioned audio-visual exposure: The effect of learner-related factors. *System* **2020**, *93*, 102303. [[CrossRef](#)]
48. Pujadas, G.; Muñoz, C. When to Switch Off Captions? In Proceedings of the EuroSLA 31st Conference, Fribourg, Switzerland, 24–27 August 2022.
49. Vanderplank, R. ‘Gist watching can only take you so far’: Attitudes, strategies and changes in behaviour in watching films with captions. *Lang. Learn. J.* **2019**, *47*, 407–423. [[CrossRef](#)]
50. Durbahn, M.; Macis, M.; Rodgers, M.; Peters, E. Lexical coverage in L2 viewing comprehension: An extension of van Zeeland and Schmitt (2012). Presented at the EuroSLA 31st Conference, Fribourg, Switzerland, 24–27 August 2022.
51. Durbahn, M.; Rodgers, M.; Peters, E. The relationship between vocabulary and viewing comprehension. *System* **2020**, *88*, 102166. [[CrossRef](#)]
52. van Zeeland, H.; Schmitt, N. Lexical coverage in L1 and L2 listening comprehension: The same or different from reading comprehension? *Appl. Linguist.* **2012**, *33*, 457–479. [[CrossRef](#)]
53. Kim, S.K.; Webb, S. Individual difference factors for L2 vocabulary. In *The Routledge Handbook of Second Language Acquisition and Individual Differences*; Li, S., Hiver, P., Papi, M., Eds.; Routledge: New York, NY, USA, 2022; pp. 282–293.
54. Lin, P.M.S.; Siyanova-Chanturia, A. Internet television for L2 vocabulary learning. In *Language Learning Beyond the Classroom*; Nunan, D., Richards, J., Eds.; Routledge: New York, NY, USA, 2015; pp. 149–158.
55. Alexiou, T. Vocabulary uptake from Peppa Pig: A case study of preschool EFL learners in Greece. In *Current Issues in Second/Foreign Language Teaching and Teacher Development: Research and Practice*; Gitsaki, C., Alexiou, T., Eds.; Cambridge Scholars Publishing: Newcastle, UK, 2015; pp. 285–301.
56. Pujadas, G.; Muñoz, C. Extensive viewing of captioned and subtitled TV series: A study of L2 vocabulary learning by adolescents. *Lang. Learn. J.* **2019**, *47*, 479–496. [[CrossRef](#)]
57. Delgiudice, M. Middle childhood: An evolutionary-developmental synthesis. In *Handbook of Life Course Health Development*; Halfon, N., Forrest, C., Lerner, R., Faustman, E., Eds.; Springer: New York, NY, USA, 2018; pp. 95–107.
58. Enever, J. *Policy and Politics in Global Primary English*; Oxford University Press: Oxford, UK, 2018.
59. Carrington, M.; Child, L. (Executive Producers); *Charlie and Lola [Animated TV Series]*; Tiger Aspects Productions: London, UK, 2005–2008.
60. Donaghy, K. Using film to teach languages in a world of screens. In *Using Film and Media in the Language Classroom: Reflections on Research-Led Teaching*; Herrero, C., Vanderchelden, I., Eds.; Multilingual Matters: London, UK, 2019; pp. 3–16.
61. Lextutor. Available online: <https://www.lex tutor.ca/vp/eng/> (accessed on 30 October 2019).
62. Santillana. *Lenguaje y Comunicación Para Cuarto Básico Serie Bicentenario*; Santillana: Santiago, Chile, 2014.
63. Cambridge Assessment English. PreA1 Starters, A1 Movers and A2 Flyers Sample Papers. Available online: <https://www.cambridgeenglish.org/exams-and-tests/movers/preparation/> (accessed on 15 October 2019).
64. Pallant, J. *SPSS Survival Manual. A Step by Step Guide to Data Analysis Using IBM SPSS*, 6th ed.; McGraw Hill Education: Berkshire, UK, 2016.
65. Puimège, E.; Peters, E. Learners’ English vocabulary knowledge prior to formal instruction: The effect of learner-related and word-related factors. *Lang. Learn.* **2019**, *69*, 943–977. [[CrossRef](#)]
66. Anthony, L.; Nation, I.S.P. Picture Vocabulary Size Test (Version 1.2.0). Available online: <http://www.laurenceanthony.net/software/pvst> (accessed on 2 May 2019).
67. Lanes, A.; Love, R.; Kalman, B.; Brenchley, M.; Pickles, M. *Updating the A2 Key and B1 Preliminary Vocabulary Lists*; Cambridge Research English Research Notes; Cambridge University Press: Cambridge, UK, 2019; pp. 1–9.
68. Avello, D. L2 Learning from Captioned-Video Viewing in Primary School Students. Ph.D. Thesis, Universitat de Barcelona, Barcelona, Spain, February 2023.
69. Pujadas, G. Language Learning through Extensive TV Viewing. A Study with Adolescent EFL Learners. Ph.D. Thesis, Universitat de Barcelona, Barcelona, Spain, November 2019.

**Disclaimer/Publisher’s Note:** The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.