MDPI

Systematic Review

Vocabulary Acquisition through Viewing Captioned and Subtitled Video: A Scoping Review and Meta-Analysis

Barry Lee Reynolds [1,2], **Yanwen Cui** [1,*], **Chian-Wen Kao** [3] **and Nathan Thomas** [4]

1. Faculty of Education, University of Macau, Macau SAR, China
2. Centre for Cognitive and Brain Sciences, University of Macau, Macau SAR, China
3. Department of Applied English, Chihlee University of Technology, New Taipei City 220305, Taiwan
4. Institute of Education, University College London, London WC1H 0AL, UK
* Correspondence: mb94886@connect.um.edu.mo

Abstract: As access to video-viewing technology has increased, so has researchers' interest in understanding how the viewing of captioned and subtitled videos can lead to effective vocabulary learning outcomes. Previously, there has been one meta-analysis on the effects of this type of video-viewing on vocabulary acquisition. However, the variables investigated and types of vocabulary knowledge analyzed were limited. To address these issues, we conducted a mixed review that combined a scoping review and meta-analysis. We identified 139 studies in major databases, of which 34 aligned with our inclusion criteria. Results from the scoping review found that researchers have assessed productive knowledge more than receptive knowledge, and knowledge of form and meaning more than knowledge of use. Participants were given TV series to view more than any other media type. Results from the meta-analysis found that viewing any type of captioned or subtitled videos had a positive effect on vocabulary acquisition. Among all the captioned and subtitled video types, viewing videos with intralingual captions had the largest effect on vocabulary learning outcomes. Furthermore, the viewing of animations had the largest effect on vocabulary learning outcomes compared with all the other types of video viewing investigated. No statistically significant difference between intentional or incidental learning conditions was found, indicating that both conditions are suitable for developing vocabulary learning through video viewing. Additional findings and implications for teaching and research are discussed.

Keywords: captioning; incidental learning; intentional learning; meta-analysis; subtitling; systematic review; video; vocabulary acquisition; vocabulary knowledge

1. Introduction

Vocabulary plays an important role in language acquisition. Developing a sizable vocabulary is necessary to use language for practical purposes [1], and multimodal input can be used to achieve this aim [2]. Video viewing, for example, has been purported to facilitate simultaneous content and vocabulary learning, especially for beginners [3]. Thus, since the 1990s, researchers have been concerned with how viewing captioned and subtitled videos might address the need to acquire a substantial amount of L2 vocabulary [4–7]. Over time, many studies have shown that viewing captioned and subtitled videos not only enhances learners' comprehension but also facilitates language acquisition [8–10], and, more specifically, vocabulary acquisition [9,11–14].

Researchers' interests in this area have led to investigations targeting numerous variables that may facilitate learning. For example, some studies compared the effect of different types of captioning and subtitling on vocabulary acquisition [14–18], intentional and incidental learning conditions [16,19,20], input medium [7,9,13,21], and vocabulary knowledge type [5,7,12,22], among others. This body of previous empirical studies has greatly contributed to the current understanding that viewing captioned and subtitled videos affects

vocabulary acquisition. However, comparing individual studies can sometimes lead to conflicting results. Thus, a synthetic review can help researchers to have a better understanding of this body of literature by providing syntheses of previous studies [23].

To our knowledge, there is just one meta-analysis that synthesized research findings on captioned videos for L2 listening and vocabulary learning [24]. Despite its inherent value, it is limited in that it neglected to investigate some important variables (see discussion below). It is also nearly a decade old. Moreover, as a meta-analysis, it neither focused on describing the overall state of knowledge nor highlighted important issues that have not yet been addressed. As such, an up-to-date mixed review can be fruitful.

Meta-analyses synthesize quantitative results of empirical studies [25], but unlike systematic or scoping reviews, they are limited in the degree of synthesis they can provide for other aspects of a field's body of work [23]. Therefore, the current mixed review first provides an overview of the characteristics of the studies that have been conducted and then updates and expands on the existing meta-analysis. In doing so, it synthesizes the scope of relevant studies, while also presenting calculations of the effect of viewing captioned and subtitled videos on vocabulary acquisition.

In order to provide broad coverage, we considered a number of potentially influential variables. For example, as the acquisition of vocabulary is not an all-or-nothing phenomenon [26], we considered whether viewing videos can lead to the acquisition of different types of vocabulary knowledge. Likewise, as video viewing most often occurs for entertainment rather than for learning purposes [3], we considered whether an intentional or incidental learning condition might affect the results of vocabulary learning from video viewing. Moreover, as different input mediums provide different motivations for viewers [27], we considered whether different mediums would affect the outcome of vocabulary acquisition from video viewing. Holistically, we synthesized the effects of different types of captioning/subtitling videos to understand potential vocabulary acquisition from different types of video viewing. As such, the current study was guided by the following research questions:

1. What types of vocabulary knowledge have been investigated in published captioning/subtitling studies?
2. What types of input media have been investigated in published captioning/subtitling studies?
3. What type of captioning/subtitling has the largest effect on vocabulary acquisition?
4. What type of input medium has the largest effect on vocabulary acquisition?
5. Does an intentional or incidental learning condition have a larger effect on vocabulary acquisition through the viewing of subtitled/captioned media?

2. Literature Review

In this section, we will briefly review relevant literature that helps to situate our research questions above. Results from our scoping review (namely, the systematic map in Supplementary Materials) and our meta-analysis build on the coverage here.

Captioning and subtitling have been shown to be effective ways of improving vocabulary acquisition for L2 learners [14,22]. In video viewing, captioning is commonly used when text functions as a service to aid hearing-impaired viewers. Thus, it usually notates sound effects and other significant audio in addition to dialogue. Captioning can also be used as a tool for language learning. Subtitling, on the other hand, is "the written translation of film dialogues appearing synchronously with the corresponding dialogues produced on the screen" [28]. It is commonly used to translate dialogue into a different language for viewers via text.

Researchers have investigated the effect of different types of captioning and subtitling on vocabulary acquisition. Intralingual subtitles are text displayed in the same language as the spoken audio dialogue. Interlingual subtitles are text displayed in a language different from the spoken audio dialogue. Bilingual subtitles are text displayed in two languages and one is usually the same as the spoken audio dialogue. Intralingual captions are the

same as intralingual subtitles with the addition of textual notations of different speakers, sound effects, music, and other dramatic audio elements. Intralingual keyword subtitles and intralingual keyword captions are the same as their intralingual subtitles or interlingual captions counterparts with the caveat of only particular words targeted for research purposes being displayed. Intralingual glossed keyword captions and intralingual glossed keyword subtitles are the same as their intralingual keyword captions and interlingual keyword subtitles counterparts but with a defining gloss provided for the displayed keywords. Intralingual full captions with highlighted keywords are the same as their intralingual caption counterpart with the addition of highlighting of particular words targeted for research purposes.

Researchers who studied the effect of subtitling [15–18] have focused on the effect of intralingual, interlingual, and bilingual subtitling. For example, Wang [18] designed a study with 80 students viewing intralingual, interlingual, and bilingual subtitling videos to compare the effect of these three subtitling types on vocabulary acquisition, finding that the effects were similar. Other researchers [15–17] compared intralingual and interlingual subtitling. Those researchers found that there were no significant differences between the effects of these two types of subtitling on vocabulary acquisition. Unfortunately, few studies have focused on the effects of captioning. Of those that did investigate captioning [7,9,12,21,29,30], most examined the effect of intralingual captioning on vocabulary acquisition. Perez at al. [9] and Teng [30] took this one step further, categorizing intralingual captioning into keyword captioning, full captioning, and full captioning with highlighted keywords. Most recently, Wang and Pellicer-Sánchez [14] recruited 112 students to view a documentary with either bilingual subtitles, captions, first language (L1) subtitles, or no subtitles. While they found the captions group incidentally acquired more word form knowledge than the bilingual subtitles group, the opposite was shown for acquiring word meaning knowledge.

In the only meta-analysis (to our knowledge) on this topic, Perez et al.'s [24] investigated the effect of different types of vocabulary knowledge on vocabulary acquisition. The researchers categorized vocabulary knowledge into vocabulary recognition and vocabulary recall. Instead of using Perez et al.'s [24] categorization, we decided to code the type of vocabulary knowledge in our meta-analysis into receptive and productive based on Nation [1], who defined receptive vocabulary knowledge as "perceiving the form of a word while listening or reading and retrieving its meaning", and productive vocabulary knowledge as "wanting to express a meaning through speaking or writing and retrieving and producing the appropriate spoken or written word form". We categorized the type of vocabulary knowledge in this way to synthesize—with more accuracy—the findings from previous studies that investigated receptive vocabulary knowledge [5,7,31], productive vocabulary knowledge [4,12,22], and both receptive and productive vocabulary knowledge [9,30]. Therefore, categorizing the type of vocabulary knowledge into receptive and productive was the most suitable approach for the current study.

The input medium is another important consideration that has been researched with great variability. For example, some researchers [5,13–15,20,30,32,33] adopted documentaries as the input medium for their studies, as documentaries have been shown to "contain more imagery in close proximity to target words than narrative TV genres" [32]. Other researchers [8,21,31,34,35] used animated content as the input medium for their studies. The rationale was that the participants in their studies were young, so animation may "entertain children-students but can also motivate better than anything else" [31]. Similarly, some researchers [7,29] chose children's television with the same rationale. Other types of input medium, such as news clips [9,36,37], movies [38,39], TV series [19,40], and instructional videos [4] were also used for various reasons. Thus, the current study collected and analyzed different types of input medium to summarize the previous results in the field of vocabulary acquisition.

Regarding learning conditions and video viewing, both incidental acquisition and intentional learning have been investigated. Researchers focusing on incidental vocabulary acquisition often argue that it is an effective way of acquiring vocabulary from context; fewer studies have investigated intentional vocabulary learning via video viewing [41]. One key criterion for truly incidental vocabulary acquisition is not informing participants that there will be a vocabulary test to follow. Some studies [9,12,22] adhered to this, while others [8,42] investigated vocabulary learning under intentional learning conditions, announcing at the beginning of the experiment that there would be a vocabulary test. These distinctions are important and have been discussed as essential differences in other meta-analyses on incidental acquisition [43]. However, Perez et al.'s [24] meta-analysis did not consider the effect of learning conditions on vocabulary acquisition. Therefore, the current study aimed to provide strong evidence that can be used to compare these two learning conditions, along with a clearer depiction of the other relevant aspects of vocabulary acquisition through viewing captioned and subtitled videos mentioned above.

3. Methodology

We conducted a sequential mixed review [44] to yield in-depth and complementary insights into vocabulary acquisition through subtitled and captioned videos. A mixed review combines two types of review, and in this study, consisted of a scoping review and meta-analysis. Mixed reviews leverage the advantages of different review types, often to summarize and interpret previous studies along with statistical data to conduct further quantitative analyses.

3.1. Scoping Reviews

Traditional narrative research reviews can be imprecise in their reported process and outcome [23]. Systematic reviews, on the other hand, tend to be "higher quality, more comprehensive, [and] less biased than other types of literature review" [45] and are better able to report transparently on broader issues than single empirical studies. The goal of a systematic review is to integrate data from different empirical studies to produce new, more holistic findings or conclusions, such as discovering relations among empirical findings [46]. This includes the development of knowledge in the area of interest but may also emphasize problems that the extant research has left unresolved [47].

Scoping reviews are one type of systematic review that focus on the extent of research activity in an area of interest to identify trends and aspects that require more attention. By mapping the research on a given topic, scoping reviews are useful before meta-analyses to determine subdomains to analyze [44]. In this study, the scoping review aimed to formulate a meaningful summary of vocabulary acquisition through viewing captioned and subtitled videos; it also provides its own implications for research and practice.

3.2. Meta-Analysis

A meta-analysis "is a statistical method for quantitatively summarizing and synthesizing data from multiple studies" [48]. It can be useful when the aim is "to consolidate similarities and clarify conflicting findings" in a pool of studies [44]. In the field of second language acquisition (SLA), it is difficult to ensure whether certain factors have positive effects on language acquisition simply based on one study's statistically significant data, because different studies may provide different results [49]. Therefore, meta-analysis is beneficial in synthesizing results from various studies to achieve a general conclusion about whether particular factors are effective.

In this study, meta-analysis is a useful tool to provide objective and quantitative results of the correlations between certain moderating variables and vocabulary acquisition. To carry out a meta-analysis, several steps should be followed [50]: (1) defining the research question, (2) searching for literature, (3) coding studies, (4) calculating an effect-size index, (5) conducting statistical analyses and interpretation, and (6) publishing the report. These steps are described in detail below.

3.3. Literature Search Procedure

To retrieve as many related journal articles as possible, we searched several prominent reference databases. The process of searching included the Education Resources Information Clearinghouse (ERIC), Scopus, Social Science Citation Index (SSCI), Arts & Humanities Citation Index (A&HCI), and Institute of Electrical and Electronics Engineers (IEEE) Xplore. ERIC, Scopus, SSCI, and A&HCI were chosen because of their superior rate of journal coverage [51], "industry standard quality assurance, and [because they] are used by the majority of academic employers and funding sources to judge research impact" [52]. Since IEEE Xplore is a research database that contains published articles mainly in the field of electrical engineering, computer science, and electronics, it was considered as a complement to check whether it could provide any related research.

A key word search was applied to select potential related articles in the above-mentioned databases. Many keywords were used to conduct searches within the selected databases and find combinations of the following search terms: [(word) OR (vocab*)] AND [(acquisition) OR (learn*)] AND [(subtitl*) OR (caption*) OR ("on screen text") OR ("onscreen text")] AND [(video*) OR (TV) OR (television*)] AND [(view*) OR (watch*)].

Within SSCI and A&HCI, the search terms to select related studies were applied by topic. Within ERIC, Scopus, and IEEE Xplore, the search terms to select related studies were applied in the title, abstract, and keywords due to the lack of a topic option in these three databases. As some databases began indexing in 1980, studies published from January 1980 to the start of the present review in February 2021 were considered.

3.4. Inclusion and Exclusion Criteria

Initially, 139 studies were identified for potential inclusion for this study through the retrieval procedure. To finalize which relevant studies were to be used, inclusion and exclusion criteria were applied to each study.

3.4.1. Inclusion Criteria

Four inclusion criteria were adopted in accessing the initial studies. First, the study must contain appropriate statistical data, such as sample sizes, mean scores, standard deviations, and so on for the calculation of effect sizes. Second, the study must investigate captioning or subtitling as modification of the original multimodal input. Fourth, vocabulary acquisition must be measured in the study. Finally, the study must be published in an academic journal.

3.4.2. Exclusion Criteria

After the selection based on the inclusion criteria, two exclusion criteria were applied to filter the studies further. First, the article was written in languages other than English (a noted limitation). Second, as captions were designed to allow deaf and hard of hearing persons to follow the action and dialogue of videos, studies with non-mainstream participants were excluded (a second noted limitation).

The application of the above-mentioned criteria resulted in the retrieval of 34 published journal articles. Figure 1 illustrates the process of identification, selection, exclusion, and inclusion of related studies.
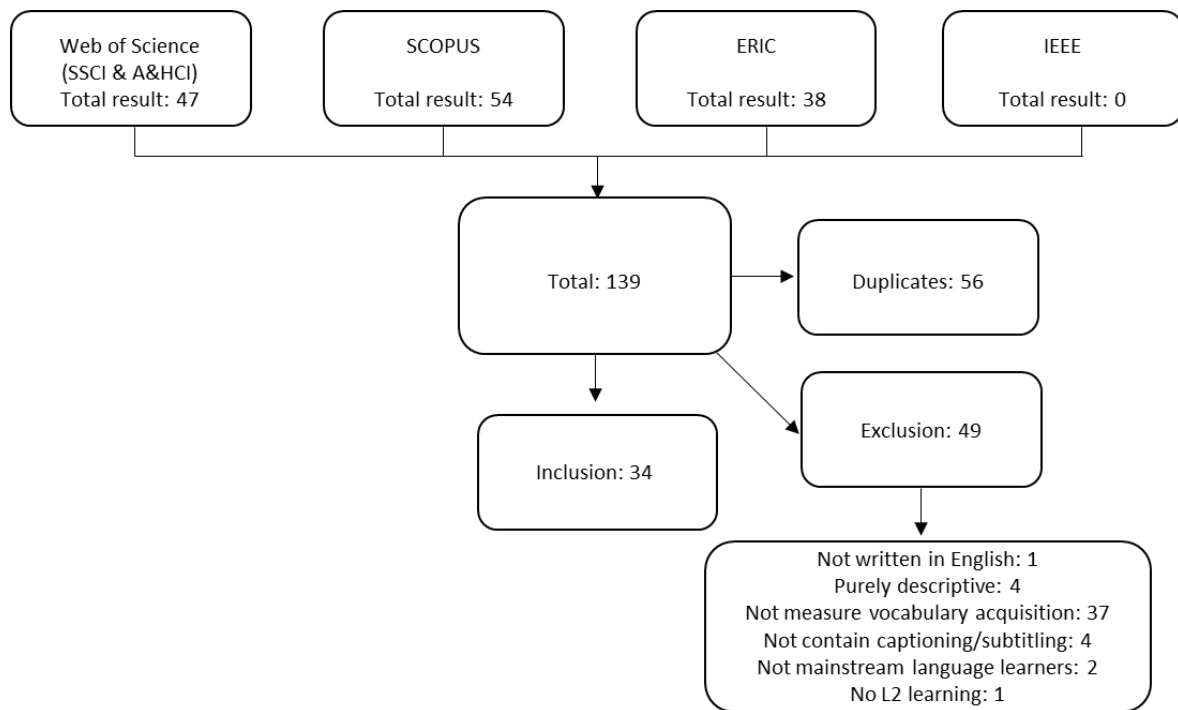
Figure 1. The Process of Study Identification and Selection.

3.5. Coding of the Studies

3.5.1. Coding Scheme

Thirty-four unique studies fulfilling the inclusion criteria were identified and then selected for this study. We coded all of the studies to identify and then classify their specific features. The moderating variables analyzed in the meta-analysis are reported in Table 1.

Table 1. Moderating Variables Coded.

Variable	Category
Type of Input Medium	1. Documentary 2. Movie 3. Instructional video 4. Children's television 5. News clip 6. Animation 7. TV series 8. Flash animation
Type of Captioning/Subtitling	1. Intralingual subtitles 2. Interlingual subtitles 3. Bilingual subtitles 4. Intralingual captions 5. Intralingual keyword captions 6. Intralingual full captions with highlighted keyword 7. Intralingual glossed keyword captions 8. Intralingual glossed keyword subtitles 9. Intralingual keyword subtitles
Learning Condition	1. Incidental 2. Intentional

Table 1. *Cont.*

Variable	Category
Assessed Vocabulary Knowledge	1. Form receptive 2. Form productive 3. Meaning receptive 4. Meaning productive 5. Use receptive 6. Use productive

3.5.2. Data Extraction

To investigate the effect of the chosen moderating variables on vocabulary acquisition, only participants' performance on vocabulary tests that assessed receptive knowledge of meaning was extracted (a noted limitation of meta-analysis). To ensure reliability, only the studies that contained both a control group and experimental group were extracted for meta-analysis, since including studies without control groups could cause erroneous results [53]. Therefore, 20 out of 34 studies were suitable for conducting the meta-analysis. We extracted the statistical data (the mean scores and standard deviations) of the vocabulary posttest results and the sample size of the control groups and experimental groups in order to calculate the effect size and conduct the further analyses. For those studies that contained more than one experimental group, we chose one particular experimental group as the representative group based on the representative moderating variable in each category of each study.

3.5.3. Effect Size Calculation

To measure the effect size, Cohen's *d* was used to calculate the results with the formula given below to investigate the effects of captioned/subtitled videos on vocabulary acquisition. Since all the primary studies included a control group and an experimental group, the following equation was used:

$$\frac{(n_1 - 1)\sigma_1 + (n_2 - 1)\sigma_2}{(n_1 - 1) + (n_2 - 1)} = \sigma_{pooled} \frac{M_1 - M_2}{\sigma_{pooled}} = d$$

In this equation, M_1 is the mean of the posttest of the experimental group, M_2 is the mean of the posttest of the control group, σ_{pooled} is the average population standard deviation, σ_1 is the standard deviation of the posttest of the experimental group, σ_2 is the standard deviation of the posttest of the control group, n_1 is the sample size of the experimental group, and n_2 is the sample size of the control group.

We used Cohen's [54] guidelines for dividing effects according to three sizes: (1) small effect size: $d = 0.20$ or $r = 0.10$; (2) medium effect size: $d = 0.50$ or $r = 0.30$; and (3) large effect size: $d = 0.80$ or $r = 0.50$.

3.5.4. Reliability

Following the advice of Cooper [23], the data were double-coded by the second author at two different times with nine months between the two rounds of coding. Intra-rater reliably reached nearly 98%. Moreover, to further increase reliability, the first author as a second coder was invited to code 10% of the papers resulting in 100% interrater agreement.

4. Results

This section describes the research features, including publication dates and participants' characteristics. After reporting the studies' basic information, we present the results of the moderating variables investigated.

4.1. Scoping Review (Research Features)

4.1.1. Published Dates

In total, 34 included studies were reviewed. All were peer-reviewed journal articles. The number of published studies increased between 1992 and 2020, and in 2019, the number of studies reached their peak (See Figure 2).
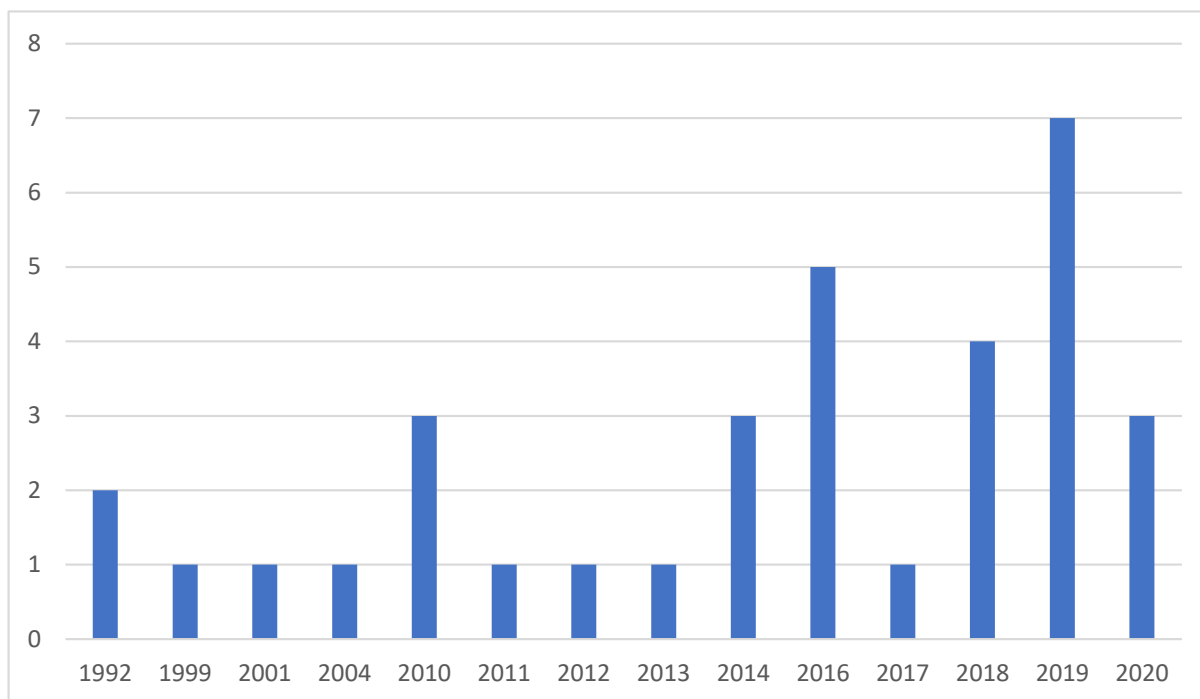

Figure 2. Published Journal Articles from 1992 to 2020.

4.1.2. Participants' Characteristics

Participants' characteristics (i.e., participants' L1, participants' L2, proficiency level, and educational level) were extracted from the studies included in the systematic review. Both Chinese and Dutch had the largest proportion of participants' L1 ($k = 7$), followed by the other languages presented in Table 2. More than 50% of the studies' participants were L2 English learners ($k = 24$), followed by French ($k = 5$), and other languages. These results are also presented in Table 2. The proficiency level, based on the Common European Framework of Reference (CEFR), was categorized into beginner, intermediate, advanced, mixed, and not reported (NR). CEFR is "a document produced by the council of Europe (CoE) to set a standard for teaching and learning English as the Second Language (ESL)" [55]. "CEFR is now internationally widespread and highly formalized into six levels: A1, A2, B1, B2, C1, and C2" [56]. The beginner level refers to A1 and A2 level in CEFR, the intermediate level refers to B1 and B2 level in CEFR, and the advanced level refers to C1 and C2 level in CEFR. Table 3 shows the participants' proficiency level, which indicates that studies with participants at the intermediate level occurred the most ($k = 16$), followed by the mixed level ($k = 8$), then beginner level ($k = 4$), and then the advanced level ($k = 2$). Four studies did not report the proficiency level (NR). Table 3 also shows participants' education level. Most studies recruited university-level learners as participants ($k = 16$), followed by primary school learners ($k = 6$), then middle school ($k = 5$), then mixed level ($k = 3$), and then vocational school ($k = 1$). Three studies did not report participants' educational level (NR).

Table 2. Learners' L1 and L2.

L1	*k*	L2	*k*
Chinese	7	English	24
Dutch	7	French	5
Arabic	2	Spanish	1
Catalan-Spanish	2	Italian	1
English	2	Dutch	1
Iranian	2	Hungarian	1
Slovenian	1	Mixed	1
Greek	1		
Japanese	1		
Italian	1		
Mainly English	1		
Spanish	1		
Romance	1		
NR	5		
Total	34	Total	34

Note. NR = Not Reported.

Table 3. Learners' Proficiency Level and Educational Level.

L2 Proficiency Level	*k*	Educational Level	*k*
Beginner	4	Primary school	6
Intermediate	16	Middle school	5
Advanced	2	Vocational school	1
Mixed	8	University	16
NR	4	Mixed	3
		NR	3
Total	34	Total	34

Note. NR = Not Reported.

4.1.3. Moderating Variables

After presenting the research features of the included studies, this section focuses on the coding of the moderating variables. For the scoping review, two moderating variables were coded and summarized: namely, vocabulary knowledge and input medium.

Vocabulary Knowledge

We coded the type of vocabulary knowledge assessed; these are presented in Table 4. After coding, the results showed that six types of vocabulary knowledge were investigated in the published captioning/subtitling studies, which were receptive knowledge of form ($k = 18$), productive knowledge of form ($k = 7$), receptive knowledge of meaning ($k = 23$), productive knowledge of meaning ($k = 16$), receptive knowledge of use ($k = 2$), and productive knowledge of use ($k = 2$). One study did not report on the type of vocabulary knowledge assessed.

Table 4. Type of Vocabulary Knowledge Assessed in Previous Studies.

Type of Vocabulary Knowledge	*k*
Receptive knowledge of meaning	23
Receptive knowledge of form	18
Productive knowledge of meaning	16
Productive knowledge of form	7

Table 4. *Cont.*

Type of Vocabulary Knowledge	*k*
Receptive knowledge of use	2
Productive knowledge of use	2
Not reported	1

Note. As some studies assessed more than one type of vocabulary knowledge, the total number equals more than 34.

Input Medium

All types of input medium in the published captioning/subtitling studies were coded and summarized. Table 5 lists the category and the number of different types of input medium, which were documentary (*k* = 7), movie (*k* = 6), instructional video (*k* = 3), children's television (*k* = 2), news clip (*k* = 4), animation (*k* = 3), TV series (*k* = 8), and flash animation (*k* = 2). The current study found that TV series were used in most of the studies, while children's television and flash animation were used least.

Table 5. Type of Input Medium.

Type of Input Medium	*k*
TV series	8
Documentary	7
Movie	6
News clip	4
Animation	3
Instructional video	3
Children's television	2
Flash animation	2
Total	35

Note. As Peters et al. (2016) used documentary and TV series as input medium, the total number equals more than 34.

4.2. Meta-Analysis

The interpretation of effect sizes is presented in this section. After presenting the overall effects, we present our investigation of the moderating variables, including the type of captioning/subtitling, the type of input medium, and learning conditions coded for further analyses.

4.2.1. Overall Captioning/Subtitling Effects

Twenty studies were included for analyses in this section. Table 6 presents overall captioning/subtitling effect sizes. Effect sizes in these 20 studies varied considerably from a large positive to a large negative effect. Large effect sizes were found in favor of captioning/subtitling under the fixed-effect model (*d* = 0.874) and the random-effects model (*d* = 0.884). Since the 95% CI did not include zero, the observed averaged effect sizes were statistically trustworthy. In addition, the test for heterogeneity indicated that the effect sizes were highly heterogeneous (I-squared: 94.783).

Table 6. Overall Captioning/Subtitling Effect Sizes (*k* = 20).

	Fixed-Effect Model	Random-Effects Model
Effect Size	0.874	0.884
Standard Error	0.054	0.241
Variance	0.003	0.058
Minimum	−1.844	−1.844

Table 6. *Cont.*

	Fixed-Effect Model	Random-Effects Model
Maximum	4.167	4.167
Upper CI	0.980	1.356
Lower CI	0.769	0.412
p-value	0.000	0.000

4.2.2. Moderating Variables

Since the effect sizes obtained in this meta-analysis were heterogeneous, it suggests that other moderating variables should be considered in investigating captioning/subtitling effectiveness [57]. Therefore, this meta-analysis further attempts to investigate the captioning/subtitling effectiveness by analyzing the following moderating variables: (1) the type of captioning/subtitling, (2) input medium, and (3) learning conditions.

Type of Captioning/Subtitling

We coded different types of captioning/subtitling into intralingual captions, intralingual subtitles, standard subtitles, and reversed subtitles. Standard subtitles refer to the displayed text being in the first language of the viewer and the audio being in the second language of the viewer while reversed subtitles refer to the displayed text being in the second language of the viewer and the audio being in the first language of the viewer [38]. Captioning/subtitling effects under the moderating variable of input modes are shown in Table 7. The effect size was large under the fixed-effect model (d = 1.814) and the random-effects model (d = 1.732) for the five studies using intralingual captions. The effect size was medium under the fixed-effect model (d = 0.594) and was large under the random-effects model (d = 0.877) for the seven studies using intralingual subtitles. The effect size was small under the fixed-effect model (d = 0.367) and was small but negative under the random-effects model (d = −0.401) for the two studies using reversed subtitles. The effect size was small under the fixed-effect model (d = 0.401) and random-effects model (d = 0.430) for the six studies using standard subtitles. Although a significant difference was found among studies using different input modes under the fixed-effect model (Q = 123.819, p = 0.000), no significant difference existed under the random-effects model (Q = 5.122, p = 0.163).

Table 7. Captioning/subtitling Effect Sizes for Different Input Modes.

Input Modes		Intralingual Captions (k = 5)		Intralingual Subtitles (k = 7)		Reversed Subtitles (k = 2)		Standard Subtitles (k = 6)	
Statistical Models		FM	RM	FM	RM	FM	RM	FM	RM
Effect Size		1.814	1.732	0.594	0.877	0.367	−0.401	0.401	0.430
Standard Error		0.101	0.638	0.087	0.239	0.253	1.404	0.100	0.233
Variance		0.010	0.408	0.008	0.057	0.064	1.971	0.010	0.054
CI (95%)	Upper	2.011	2.983	0.764	1.346	0.863	2.351	0.597	0.886
	Lower	1.616	0.480	0.424	0.408	−0.128	−3.152	0.204	−0.027
p-value		0.000	0.007	0.000	0.000	0.147	0.775	0.000	0.065

Note. CI: Confidence Interval; FM: Fixed-effect model; RM: Random-effects model.

Type of Input Medium

For the meta-analysis, we recoded the type of input medium in order to further calculate and analyze the effect of the type of input medium on vocabulary acquisition. Captioning/subtitling effects under the moderating variable of input medium are shown in Table 8. The effect size was large under the fixed-effect model ($d = 1.674$) and the random-effects model ($d = 1.652$) for the four studies using animation. The effect size was large under the fixed-effect model ($d = 0.824$) and was large under the random-effects model ($d = 0.851$) for the three studies using documentaries. The effect size was large under the fixed-effect model ($d = 0.949$) and the random-effects model ($d = 0.936$) for the four studies using instructional videos. The effect size was negligible under the fixed-effect model ($d = -0.022$) and was negatively medium under the random-effects model ($d = -0.775$) for the two studies using movies. The effect size was large under the fixed-effect model ($d = 1.000$) and the random-effects model ($d = 1.126$) for the four studies using news clips. The effect size small under the fixed-effect model ($d = 0.251$) and the random-effects model ($d = 0.453$) for the three studies using TV series. Although a significant difference was found among studies using different input media under the fixed-effect model ($Q = 88.610$, $p = 0.000$), no significant difference existed under the random-effects model ($Q = 4.396$, $p = 0.494$).

Learning Conditions

We divided learning conditions into intentional learning and incidental learning. Captioning/subtitling effects under the moderating variable of learning conditions are shown in Table 9. The effect size was large under the fixed-effect model ($d = 0.897$) and the random-effects model ($d = 0.836$) for the 14 studies under incidental conditions. The effect size was medium under the fixed-effect model ($d = 0.798$) and was large under the random-effects model ($d = 0.905$) for the six studies under intentional learning conditions. No significant difference was evident under the fixed-effect model ($Q = 0.602$, $p = 0.438$) and the random-effects model ($Q = 0.033$, $p = 0.856$).

Table 8. Captioning/subtitling Effect Sizes for Different Input Medium.

Input Modes		Animation (k = 4)		Documentary (k = 3)		Instructional Video (k = 4)		Movie (k = 2)		News Clip (k = 4)		TV Series (k = 3)	
Statistical Models		**FM**	**RM**	**FM**	**RM**	**FM**	**RM**	**FM**	**RM**	**FM**	**RM**	**FM**	**RM**
Effect Size		1.674	1.652	0.824	0.851	0.949	0.936	−0.022	−0.775	1.000	1.126	0.251	0.453
Standard Error		0.124	1.005	0.105	0.501	0.121	0.208	0.165	0.999	0.150	0.329	0.157	0.618
Variance		0.015	1.010	0.011	0.251	0.015	0.043	0.027	0.999	0.022	0.108	0.025	0.382
CI (95%)	Upper	1.916	3.621	1.029	1.833	1.186	1.343	0.301	1.184	1.294	1.771	0.558	1.665
	Lower	1.431	−0.318	0.618	−0.131	0.711	0.529	−0.344	−2.734	0.707	0.482	−0.057	−0.758
p-value		0.000	0.100	0.000	0.089	0.000	0.000	0.896	0.438	0.000	0.001	0.110	0.463

Note. CI: Confidence Interval; FM: Fixed-effect model; RM: Random-effects model.

Table 9. Captioning/subtitling Effect Sizes for Different Learning Conditions.

Input Modes		Incidental Learning ($k = 14$)		Intentional Learning ($k = 6$)	
Statistical Models		**FM**	**RM**	**FM**	**RM**
Effect Size		0.897	0.836	0.798	0.905
Standard Error		0.061	0.324	0.112	0.203
Variance		0.004	0.105	0.012	0.041
CI (95%)	Upper	1.017	1.471	1.017	1.304
	Lower	0.777	0.200	0.580	0.507
p-value		0.000	0.010	0.000	0.000

Note. CI: Confidence Interval; FM: Fixed-effect model; RM: Random-effects model.

5. Discussion

5.1. The Overall Effectiveness of Captioned and Subtitled Video Viewing on Vocabulary Acquisition

The primary aim of this review was to present the overall effect of captioned and subtitled video viewing on vocabulary acquisition. The large effect sizes found in favor of captioning/subtitling under both the fixed-effect model ($d = 0.874$) and the random-effects model ($d = 0.884$) indicated that viewing captioned and subtitled videos has a large positive effect on vocabulary acquisition.

Classic SLA theories such as the input hypothesis and dual coding theory provide theoretical explanations for why captioned and subtitled videos are effective multimodal aids to vocabulary acquisition. The input hypothesis put forward by Krashen [58] indicated that learners can acquire vocabulary knowledge with ample comprehensible input. Therefore, learners who have access to more comprehensible input can acquire vocabulary knowledge more effectively. For literate learners, captions/subtitles likely make input more comprehensible. According to dual-coding theory [59], learners will acquire vocabulary knowledge more effectively if it is presented through a dual rather than single code. When learners are immersed in multimedia environments that expose them to multimodal input (e.g., textual, graphic, and auditory resources), they are likely to retain greater vocabulary knowledge [60]. When captioned and subtitled videos are presented, the audio of the videos combine with the captions/subtitles to increase the comprehensibility of the language input [60]. These classic theories help to explain the importance of captioned and subtitled videos as not only a source of comprehensible input but also as a route to receiving a dual mode of verbal and visual information that results in effective vocabulary acquisition and retention.

5.2. Assessed Vocabulary Knowledge

Nation [1] categorized vocabulary knowledge into six types (see Table 4). Perez et al.'s, [24] meta-analysis on the effect of viewing captioned videos coded vocabulary knowledge into *recognition* and *recall*. The terms *recall*, *recognition*, *receptive,* and *productive* were also used in the synthesized primary studies to refer to vocabulary assessments but without clearly stating the type of vocabulary knowledge assessed by them. Confusing labels for different types of vocabulary knowledge and assessments has been noted by previous vocabulary researchers [61]. The current study addressed these limitations through the more nuanced investigation of the six types of vocabulary knowledge clearly defined by [1]. We found that receptive vocabulary knowledge was assessed more than productive knowledge. Moreover, the vocabulary knowledge of form and meaning were assessed more than the vocabulary knowledge of use.

There are differences between receptive and productive vocabulary knowledge that may cause researchers to be more or less inclined to assess these knowledge types. "Receptive vocabulary involves perceiving the form of a word while listening or reading and

retrieving its meaning . . . productive vocabulary involves the desire to express a meaning through speaking or written word form" [1]. A learner will usually acquire "vocabulary words receptively first and only after intentional learning they become available for their productive use" [62]. In other words, there is the potential for a continuum between receptive to productive knowledge that grows over time after repeated exposures [63]. If learners are only exposed to new words for a short time through video viewing, they may only be able to develop receptive knowledge. Participants recruited for the reviewed primary studies may not have been given sufficient time for productive knowledge development [5,9,18,31]. This could explain the reason researchers opted to assess receptive vocabulary knowledge more than productive vocabulary knowledge.

Nation [1] explains that vocabulary knowledge of use involves grammar, collocation, and constraints which require both implicit and explicit learning. Vocabulary knowledge of use is difficult for learners to acquire and complex for researchers to assess. To assess vocabulary knowledge of use, researchers need to design assessments that require participants to write sentences or produce the words in oral production tasks. Therefore, practical reasons may have dictated the types of vocabulary knowledge assessed: it is quicker and more straightforward to administer and mark receptive vocabulary knowledge assessments. Moreover, six primary studies included young learners as participants [5,29–32,35]. "Young learners are notoriously poor test taker[s] . . . The younger the child being evaluated, assessed, or tested, the more errors are made" [64]. In addition, as much of the existing research had already assessed receptive knowledge of form and meaning, researchers may have also selected these types of knowledge so comparisons could be made.

5.3. Input Medium

Researchers choose different mediums to suit different participant groups. TV series were used the most, while children's television and flash animations were used the least. Most of the primary studies recruited university students as participants, so TV series, movies, and documentaries might have been found more suitable [11,13,18,20,22,33,39,65–67]. Unlike TV series that have been found to be suitable for participants of different ages [15,16,66,67], children's television and flash animations may have been shown to younger participants as children consider these media types more entertaining [31].

Among the six media types, a large effect was found for animations even though they were used the least by researchers. The studies involving animations recruited primary and middle school students as participants [8,31,34,35]. Animations have been shown to "grasp children's attention quickly . . . [and] sharpe[n] their observations" [68]. While there might have been an entertaining element that led to this large effect, it is likely that after drawing in and maintaining their attention, the animations were able to help the learners understand better.

5.4. Captioning/Subtitling Type

Bird and Williams [69] and Schmidt [70] claimed that intralingual captions/subtitles are more effective than interlingual captions/subtitles because "word boundaries are clear and there are no accent variations, [so] language learners comprehend and learn language to a greater extent" [71]. The present meta-analysis supports these claims as we found the intralingual condition more effective than the reversed and standard (interlingual condition). The intralingual condition helps learners "to recognize the words that are being spoken" [72], because the pronunciation of words in the audio and the spelling of words on the screen are in the same language. This could assist learners in making the form-meaning link, an important initial step necessary when learning new words [1]. The present meta-analysis also found intralingual captions more effective than intralingual subtitles. As captions provide more information and create a more comprehensible environment for learners, this could further scaffold the learners' comprehension of the video content and thereby further increase the likelihood of vocabulary learning. Although researcher involvement in enhancing the input media would be increased, it may also be fruitful to

consider the implications of various types of textually enhanced subtitles/captions (e.g., bold, underlined, and text in different colors), as this continues to be a promising area of research in studies on L2 development [73,74].

5.5. Intentional and Incidental Learning Conditions

The meta-analysis failed to find a significant difference between intentional or incidental learning of vocabulary from viewing captioned and subtitled videos. However, both intentional and incidental conditions were found to be effective, showing medium-to-large effect sizes. Although the present meta-analysis operationalized incidental learning by whether or not the participants were informed of an upcoming vocabulary test [75], most learners will "preserve a certain amount of attentional resources for processing unknown words" [76] in videos even though they may have been placed in an incidental learning condition. When learners begin to notice an unknown word in the on-screen text, the distinction between intentional and incidental learning conditions begins to blur. Similarly, placing learners in an intentional learning condition while viewing videos may not result in an increase in vocabulary acquisition as the verbal and visual information presented cannot be controlled or reviewed. Unlike static text on a page that can be fixated on for long periods of time or reviewed through regressions, the text in captions and subtitles will disappear and be replaced immediately as the video plays.

6. Conclusions

The present mixed synthesis found intralingual captions and animations the most effective. These results have practical implications for instructors that opt to incorporate video viewing into their language classrooms. Teachers should provide intralingual captioned videos to learners if the video viewing is for the purpose of vocabulary learning. For young learners, animations should be provided. For learners in a foreign-language context, we also feel that teachers should encourage video viewing beyond the classroom as extensive viewing has been purported to compensate for the lack of language input outside the classroom [3,16]. For example, if teachers play one episode of an animation series for learners inside the classroom, teachers may suggest learners continue viewing the remaining episodes outside of the classroom. This may be especially useful during school breaks such as summer holidays when language input in foreign language settings is often further limited.

The results of this study have allowed us to identify several issues that deserve future research attention. As most of the primary studies assessed receptive vocabulary knowledge, our current understanding of productive vocabulary knowledge acquisition from captioned and subtitled video viewing is rather limited. Given the continued attention to input processing in SLA, it would be interesting to see future researchers focus investigations and critical discussions on how language input through viewing captioned and subtitled video can more or less lead to productive vocabulary knowledge growth. Similarly, few studies assessed vocabulary knowledge of use, adding little to our understanding on how viewing captioned and subtitled videos influences learners' ability to use new words productively that have been encountered receptively in video input. Vocabulary knowledge of use involves grammar, collocation, and constraints on use, which requires both implicit and explicit learning [1], making knowledge of use complex for researchers to assess and difficult for learners to master. Researchers may address this gap in the literature by designing relevant studies that aim to investigate the acquisition of knowledge of vocabulary use with participants that are exposed to target words over a longer period of time.

Although animations were found to be effective at inducing vocabulary learning, the majority of those studies recruited young learners as participants. Future studies should recruit participants of different ages and language proficiencies to examine whether these moderate the positive effect shown for viewing animations.

While no statistical difference was found between vocabulary learning under intentional and incidental learning conditions, the number of studies conducted under an incidental condition were quite uneven and reduced statistical power. We feel there is reason to call for an empirical study that determines experimentally whether these two conditions are likely to induce different learning outcomes in similar ways to text-based language input.

While the present mixed review provides an understanding of the potential vocabulary acquisition that results from different types of video viewing, it is not without limitations. The reader should bear in mind that the meta-analysis was completed by extracting vocabulary learning outcomes for only one type of knowledge. As only a single effect size could be extracted from each primary study, this limited our investigation to the aspect of vocabulary knowledge measured by most of the primary studies—receptive knowledge of meaning. As few studies measured retention with delayed posttests, data extraction was limited to posttest scores. Another potential limitation is that most of the primary studies recruited students as research participants. We consider it best practice to delimit our findings to the acquisition of receptive knowledge of meaning by language learners that are pursuing some level of formal education.

References

1. Nation, I.S.P. *Learning Vocabulary in Another Language*; Cambridge University Press: Cambridge, UK, 2013.
2. Pellicer-Sánchez, A. Multimodal reading and second language learning. *ITL Int. J. Appl. Linguist.* **2022**, *173*, 2–17. [CrossRef]
3. Webb, S. Extensive viewing: Language learning through watching television. In *Language Learning Beyond the Classroom*; Nunan, D., Richard, J.C., Eds.; Routledge: London, UK, 2015; pp. 159–168.
4. Danan, M. Reversed Subtitling and Dual Coding Theory: New Directions for Foreign Language Instruction. *Lang. Learn.* **1992**, *42*, 497–527. [CrossRef]
5. Koolstra, C.M.; Beentjes, J.W.J. Children's vocabulary acquisition in a foreign language through watching subtitled television programs at home. *Educ. Technol. Res. Dev.* **1999**, *47*, 51–60. [CrossRef]
6. Koskinen, P.S.; Wilson, R.M.; Gambrell, L.B.; Neuman, S.B. Captioned video and vocabulary learning: An innovative practice in literacy instruction. *Read. Teach.* **1993**, *47*, 36–43. Available online: http://www.jstor.org/stable/20201190 (accessed on 1 February 2021).
7. Neuman, S.B.; Koskinen, P. Captioned Television as Comprehensible Input: Effects of Incidental Word Learning from Context for Language Minority Students. *Read. Res. Q.* **1992**, *27*, 94–106. [CrossRef]
8. Lwo, L.; Lin, M.C.-T. The effects of captions in teenagers' multimedia L2 learning. *ReCALL* **2012**, *24*, 188–208. [CrossRef]
9. Perez, M.M.; Peters, E.; Clarebout, G.; Desmet, P. Effects of captioning on video comprehension and incidental vocabulary learning. *Lang. Learn. Technol.* **2014**, *18*, 118–141.
10. Yuksel, D.; Tanriverdi, B. Effect of watching captioned movie clip on vocabulary development of EFL learners. *Turk. Online J. Educ. Technol.* **2009**, *8*, 48–54. Available online: http://www.tojet.net/articles/v8i2/824.pdf (accessed on 1 February 2021).

11. Alharthi, T. Can Adults Learn Vocabulary through Watching Subtitled Movies? An Experimental Corpus-Based Approach. *Int. J. Engl. Lang. Lit. Stud.* **2020**, *9*, 219–230. [CrossRef]

12. Harji, M.B.; Alavi, Z.K.; Letchumanan, K. Captioned Instructional Video: Effects on Content Comprehension, Vocabulary Acquisition and Language Proficiency. *Engl. Lang. Teach.* **2014**, *7*, 1–16. [CrossRef]

13. Fang, F.; Zhang, Y.; Fang, Y. A Comparative Study of the Effect of Bilingual Subtitles and English Subtitles on College English Teaching. *Rev. Cercet. Interv. Sociala* **2019**, *66*, 59–74. [CrossRef]

14. Wang, A.; Pellicer-Sánchez, A. Incidental Vocabulary Learning from Bilingual Subtitled Viewing: An Eye-Tracking Study. *Lang. Learn.* **2022**, *72*, 765–805. [CrossRef]

15. Peters, E.; Heynen, E.; Puimège, E. Learning vocabulary through audiovisual input: The differential effect of L1 subtitles and captions. *System* **2016**, *63*, 134–148. [CrossRef]

16. 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]

17. Stewart, M.A.; Pertusa, I. Gains to Language Learners from Viewing Target Language Closed-Captioned Films. *Foreign Lang. Ann.* **2004**, *37*, 438–442. [CrossRef]

18. Wang, Y. Effects of L1/L2 Captioned TV Programs on Students' Vocabulary Learning and Comprehension. *CALICO J.* **2019**, *36*, 204–224. [CrossRef]

19. Aidinlou, N.A. Short-term and Long-term Retention of Vocabulary through Authentic Subtitled Videos. *Adv. Lang. Lit. Stud.* **2016**, *7*, 14–22. [CrossRef]

20. Perez, M.M. Pre-learning vocabulary before viewing captioned video: An eye-tracking study. *Lang. Learn. J.* **2019**, *47*, 460–478. [CrossRef]

21. Mohsen, M.A. Effects of help options in a multimedia listening environment on L2 vocabulary acquisition. *Comput. Assist. Lang. Learn.* **2016**, *29*, 1220–1237. [CrossRef]

22. Ashcroft, R.J.; Garner, J.; Hadingham, O. Incidental vocabulary learning through watching movies. *Aust. J. Appl. Linguist.* **2018**, *1*, 135–147. [CrossRef]

23. Cooper, H. *Research Synthesis and Meta-Analysis: A Step-by-Step Approach*, 4th ed.; Sage: Thousand Oaks, CA, USA, 2017. [CrossRef]

24. Perez, M.M.; Noortgate, W.V.D.; Desmet, P. Captioned video for L2 listening and vocabulary learning: A meta-analysis. *System* **2013**, *41*, 720–739. [CrossRef]

25. Hak, T.; Van Rhee, H.; Suurmond, R. *How to Interpret Results of Meta-Analysis*; SSRN: Rochester, NY, USA, 2016. [CrossRef]

26. Laufer, B. The Development of Passive and Active Vocabulary in a Second Language: Same or Different? *Appl. Linguist.* **1998**, *19*, 255–271. [CrossRef]

27. Hanson, G.; Haridakis, P. YouTube Users Watching and Sharing the News: A Uses and Gratifications Approach. *J. Electron. Publ.* **2008**, *11*, 36–38. [CrossRef]

28. Gorjian, B. The Effect of Movie Subtitling on Incidental Vocabulary Learning among EFL Learners. *Int. J. Asian Soc. Sci.* **2014**, *4*, 1013–1026.

29. Linebarger, D.L. Learning to read from television: The effects of using captions and narration. *J. Educ. Psychol.* **2001**, *93*, 288–298. [CrossRef]

30. Teng, F. Incidental vocabulary learning for primary school students: The effects of L2 caption type and word exposure frequency. *Aust. Educ. Res.* **2018**, *46*, 113–136. [CrossRef]

31. Ina, L. Incidental foreign-language acquisition by children watching subtitled television programs. *Turk. Online J. Educ. Technol.* **2014**, *13*, 81–87. Available online: http://www.tojet.net/articles/v13i4/1349.pdf (accessed on 1 February 2021).

32. Peters, E. The Effect of Imagery and On-Screen Text on Foreign Language Vocabulary Learning from Audiovisual Input. *TESOL Q.* **2019**, *53*, 1008–1032. [CrossRef]

33. Winke, P.; Gass, S.; Sydorenko, T. The effects of captioning videos used for foreign language listening activities. *Lang. Learn. Technol.* **2010**, *14*, 65–86.

34. Chen, Y.-R.; Liu, Y.-T.; Todd, A.G. Transient but Effective? Captioning and Adolescent EFL Learners' Spoken Vocabulary Acquisition. *Engl. Teach. Learn.* **2018**, *42*, 25–56. [CrossRef]

35. Teng, F. The effects of video caption types and advance organizers on incidental L2 collocation learning. *Comput. Educ.* **2019**, *142*, 103655. [CrossRef]

36. Fievez, I.; Perez, M.M.; Cornillie, F.; Desmet, P. Vocabulary Learning Through Viewing Captioned or Subtitled Videos and the Role of Learner- and Word-Related Factors. *CALICO J.* **2020**, *37*, 233–253. [CrossRef]

37. Sirmandi, E.H.; Sardareh, S.A. The effect of BBC world clips with and without subtitles on intermediate EFL learners' vocabulary development. *Malays. Online J. Educ. Sci.* **2016**, *4*, 61–69. Available online: https://mojes.um.edu.my/index.php/MOJES/article/view/12674/8159 (accessed on 1 February 2021).

38. Bisson, M.-J.; VAN Heuven, W.J.B.; Conklin, K.; Tunney, R.J. Processing of native and foreign language subtitles in films: An eye tracking study. *Appl. Psycholinguist.* **2012**, *35*, 399–418. [CrossRef]

39. Perego, E.; Del Missier, F.; Porta, M.; Mosconi, M. The Cognitive Effectiveness of Subtitle Processing. *Media Psychol.* **2010**, *13*, 243–272. [CrossRef]

40. Birulés-Muntané, J.; Soto-Faraco, S. Watching Subtitled Films Can Help Learning Foreign Languages. *PLoS ONE* **2016**, *11*, e0158409. [CrossRef]

41. Thomas, N. Incidental L2 vocabulary learning: Recent developments and implications for future research. *Read. Foreign Lang.* **2020**, *32*, 49–60.
42. Suárez, M.D.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*, 497–517. [CrossRef]
43. De Vos, J.F.; Schriefers, H.; Nivard, M.; Lemhöfer, K. A Meta-Analysis and Meta-Regression of Incidental Second Language Word Learning from Spoken Input. *Lang. Learn.* **2018**, *68*, 906–941. [CrossRef]
44. Chong, S.W.; Plonsky, L. *A Typology of Secondary Research in Applied Linguistics*; OSF Preprint: Charlottesville, VA, USA, 2021. [CrossRef]
45. Siddaway, A.P.; Wood, A.M.; Hedges, L.V. How to Do a Systematic Review: A Best Practice Guide for Conducting and Reporting Narrative Reviews, Meta-Analyses, and Meta-Syntheses. *Annu. Rev. Psychol.* **2018**, *70*, 747–770. [CrossRef]
46. Pollock, A.; Berge, E. How to do a systematic review. *Int. J. Stroke* **2017**, *13*, 138–156. [CrossRef] [PubMed]
47. Thomas, N.; Bowen, N.E.J.A.; Reynolds, B.L.; Osment, C.; Pun, J.K.H.; Mikolajewska, A. A Systematic Review of the Core Components of Language Learning Strategy Research in Taiwan. *Engl. Teach. Learn.* **2021**, *45*, 355–374. [CrossRef]
48. In'nami, Y.; Koizumi, R.; Tomita, Y. Meta-analysis in applied linguistics. In *The Routledge Handbook of Research Methods in Applied Linguistics*; McKinley, J., Rose, H., Eds.; Routledge: London, UK, 2020; pp. 240–252.
49. Norris, J.M.; Ortega, L. The value and practice of research synthesis for language learning and teaching. In *Synthesizing Research on Language Learning and Teaching*; Norris, J.M., Ortega, L., Eds.; Benjamins: Amsterdam, The Netherlands, 2006; pp. 3–52.
50. Sánchez-Meca, J.; Marín-Martínez, F. Meta-analysis in psychological research. *Int. J. Psychol. Res.* **2010**, *3*, 150–162. [CrossRef]
51. In'Nami, Y.; Koizumi, R. Database Selection Guidelines for Meta-Analysis in Applied Linguistics. *TESOL Q.* **2010**, *44*, 169–184. [CrossRef]
52. Thomas, N.; Bowen, N.E.J.A.; Rose, H. A diachronic analysis of explicit definitions and implicit conceptualizations of language learning strategies. *System* **2021**, *103*, 102619. [CrossRef]
53. Hunter, J.E.; Jensen, J.L.; Rodgers, R. The Control Group and Meta-Analysis. *J. Methods Meas. Soc. Sci.* **2014**, *5*, 3–21. [CrossRef]
54. Cohen, J. *Statistical Power Analysis for the Behavioral Sciences*; Lawrence Erlbaum Associates: Mahwah, NJ, USA, 1998.
55. Abidin, N.Z.; Hashim, H. Common European Framework of Reference (CEFR): A Review on Teachers' Perception & Plurilingualism. *Creative Educ.* **2021**, *12*, 727–736. [CrossRef]
56. Kihlstedt, M. Foreign Language (FL) Teaching and Learning in Primary Schools in Europe: Beliefs and Realities. *Emerg. Trends Educ.* **2019**, *2*, 71–96. [CrossRef]
57. Borenstein, M.; Hedges, L.V.; Higgins, J.P.; Rothstein, H.R. *Introduction to Meta-Analysis*, 2nd ed.; John Wiley & Sons: Hoboken, NJ, USA, 2009.
58. Krashen, S. *The Input Hypothesis: Issues and Implications*; Laredo Publishing Company: Englewood, NJ, USA, 1985.
59. Mezei, A.; Paivio, A. Imagery and Verbal Processes. *Leonardo* **1971**, *5*, 359. [CrossRef]
60. Kanellopoulou, C.; Kermanidis, K.L.; Giannakoulopoulos, A. The Dual-Coding and Multimedia Learning Theories: Film Subtitles as a Vocabulary Teaching Tool. *Educ. Sci.* **2019**, *9*, 210. [CrossRef]
61. Schmitt, N. *Researching Vocabulary: A Vocabulary Research Manual*; Palgrave Macmillan: London, UK, 2010.
62. Zhou, S. Comparing Receptive and Productive Academic Vocabulary Knowledge of Chinese EFL Learners. *Asian Soc. Sci.* **2010**, *6*, 14–19. [CrossRef]
63. Pignot-Shahov, V. Measuring L2 receptive and productive vocabulary knowledge. *Lang. Stud. Work. Pap.* **2012**, *4*, 37–45. Available online: https://www.reading.ac.uk/elal/-/media/project/uor-main/schools-departments/elal/lswp/lswp-4/elal_lswp_vol_4_pignot_shahov.pdf?la=en&hash=0FC24F4853501697EA680CB52FC24C87 (accessed on 1 February 2021).
64. Katz, L.G. *A Developmental Approach to Assessment of Young Children*; University of Illinois: Urbana, IL, USA, 1997.
65. Čepon, S. Effective Use of the Media: Video in the Foreign Language Classroom. *Medijska Istraž. Znan. Stručni Čas. Novin. Medije* **2013**, *19*, 83–104.
66. Hsu, H.-T. Incidental professional vocabulary scquisition of EFL business learners: Effect of captioned video with glosses as a multimedia annotation. *JALT CALL J.* **2018**, *14*, 119–142. [CrossRef]
67. Sinyashina, E. Watching Captioned Authentic Videos for Incidental Vocabulary Learning: Is It Effective? *Nord. J. Engl. Stud.* **2020**, *19*, 28–64. [CrossRef]
68. Ghilzai, S.A.; Alam, R.; Ahmad, Z.; Shaukat, A.; Noor, S.S. Impact of cartoon programs on children's language and behavior. *Insights Lang. Soc. Cult.* **2017**, *2*, 104–126.
69. 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]
70. Schmidt, C. Same Language Subtitling on Television: A Tool for Promoting Literacy Retention in India? Master's Thesis, Stanford University, Stanford, CA, USA, 2007.
71. Zarei, A.A.; Rashvand, Z. The Effect of Interlingual and Intralingual, Verbatim and Nonverbatim Subtitles on L2 Vocabulary Comprehension and Production. *J. Lang. Teach. Res.* **2011**, *2*, 618–625. [CrossRef]
72. Baranowska, K. Learning most with least effort: Subtitles and cognitive load. *ELT J.* **2020**, *74*, 105–115. [CrossRef]
73. Chung, Y.; Révész, A. Investigating the effect of textual enhancement in post-reading tasks on grammatical development by child language learners. *Lang. Teach. Res.* **2021**, *7*, 13621688211005068. [CrossRef]

74. Révész, A.; Bunting, L.; Florea, A.; Gilabert, R.; Segerstad, Y.H.A.; Mihu, I.P.; Parry, C.; Benton, L.; Vasalou, A. The Effects of Multiple-Exposure Textual Enhancement on Child L2 Learners' Development in Derivational Morphology: A Multi-Site Study. *TESOL Q.* **2021**, *55*, 901–930. [CrossRef]

75. Hulstijn, J.H. Incidental and Intentional Learning. In *The Handbook of Second Language Acquisition*; Doughty, C.J., Long, M.H., Eds.; Blackwell: Hoboken, NJ, USA, 2003; pp. 349–381. [CrossRef]

76. Perez, M.M.; Peters, E.; Desmet, P. Vocabulary learning through viewing video: The effect of two enhancement techniques. *Comput. Assist. Lang. Learn.* **2018**, *31*, 1–26. [CrossRef]

77. Harji, M.B.; Woods, P.C.; Alavi, Z.K. The Effect of Viewing Subtitled Videos on Vocabulary Learning. *J. Coll. Teach. Learn.* **2010**, *7*, 37–42. [CrossRef]

78. Baddeley, A.D. *Working Memory*; Oxford University Press: Oxford, UK, 1986.

79. Fletcher, J.D.; Tobias, S. The multimedia principle. In *The Cambridge handbook of multimedia learning*; Mayer, R.E., Ed.; Cambridge University Press: Cambridge, UK, 2005; pp. 117–133. [CrossRef]

80. Hsia, J. The information capacity of modality and channel performance. *Audio Vis. Commun. Rev.* **1971**, *19*, 51–75. [CrossRef]

81. Mayer, R. Multimedia learning: Are we asking the right questions? *Educ. Psychol.* **1997**, *32*, 1–19. [CrossRef]

82. Mayer, R.E. *Multimedia Learning*; Cambridge University Press: Cambridge, UK, 2001. [CrossRef]

83. Paivio, A. *Imagery and Verbal Processes*; Psychology Press: London, UK, 1979. [CrossRef]

84. Wang, Y.-C. Learning L2 vocabulary with American TV drama from the learner's perspective. *Engl. Lang. Teach.* **2012**, *5*, 217–225. [CrossRef]