Computer Assisted Language Instruction Consortium

CALICO JOURNAL (ONLINE) ISSN 2056–9017

Article

Vocabulary Learning Through Viewing Captioned or Subtitled Videos and the Role of Learner- and Word-Related Factors

Isabeau Fievez[1], Maribel Montero Perez[2], Frederik Cornillie[3], and Piet Desmet[4]

Abstract

This study investigates incidental vocabulary learning through captioned or subtitled videos and examines whether and how different learner- (prior vocabulary knowledge) and word-related factors (i.e., frequency of occurrence, cognateness, and imagery) influence learning gains from watching videos. Low-intermediate Dutch-speaking learners of French (N=86) took part in a four week intervention program. They were assigned to a subtitles group, a captions group, or a control group (who only took the tests). Vocabulary learning was measured by means of form and meaning recognition, as well as meaning recall tests.

Results revealed that participants learned approximately 15% of the vocabulary they could have learned. Both treatment groups outperformed the control group in the meaning recognition test, but only the captions group outperformed the control group in the meaning recall test. Learning gains were mediated by cognateness with significantly higher odds to recall and recognize a cognate on the posttest than a noncognate. Frequency of occurrence and prior vocabulary knowledge had a positive effect on L2 learners' ability to recall and recognize the meaning of the target words. A positive relationship was also

Affiliations

[1]Department of Linguistics, KU Leuven, Belgium and ITEC, imec research group at KU Leuven, Belgium.
email: isabeau.fievez@kuleuven.be

[2]Department of Linguistics, Ghent University, Belgium.
email: maribel.monteroperez@ugent.be

[3]Department of Linguistics, KU Leuven, Belgium and ITEC, imec research group at KU Leuven, Belgium.
email: frederik.cornillie@kuleuven.be

[4]Department of Linguistics, KU Leuven, Belgium and ITEC, imec research group at KU Leuven, Belgium.
email: piet.desmet@kuleuven.be

CALICO JOURNAL VOL 37.3 2020 233–253 https://doi.org/10.1558/cj.39370

equinox www.equinoxpub.com

equinoxonline

found between target words that were visually represented in the video and learners' meaning recall scores for those words.

Keywords: incidental vocabulary learning; videos; captions; frequency of occurrence; imagery.

1. Introduction

In the past decades, the research fields of second language acquisition (SLA) and computer-assisted language learning (CALL) have shown an increased interest in the potential of audio-visual input for incidental vocabulary learning (Montero Perez & Rodgers, 2019). Through the growing offer of digital platforms for video distribution, such as YouTube and Netflix, second language (L2) learners now have a plethora of opportunities to engage with authentic audio-visual L2 input, primarily outside formal learning contexts but also in the classroom.

Numerous studies have investigated whether watching videos with subtitles or captions—on-screen text in the L1 or the L2 respectively—results in incidental vocabulary learning gains (e.g., Peters, Heynen, & Puimège, 2016). Yet, findings with regard to the differential impact of L1 and L2 subtitles on various aspects of word knowledge remain inconclusive.

In addition, most studies on incidental vocabulary learning through viewing have made use of short clips in single treatment sessions (e.g., Montero Perez, Peters, & Desmet, 2018), whereas one of the indisputable tenets of L2 incidental language learning theories is that exposure to the language is fundamental (Nation, 2013). In addition, there is beginning evidence that several learner- and word-related factors influence vocabulary gains through viewing (e.g., Peters, 2019), but little is known about their role in studies using multiple viewing sessions.

Therefore, this study researches the effects of watching subtitled or captioned videos on incidental vocabulary learning in a four-week classroom intervention study. It also investigates the role of learner- and word-related factors. Another important feature of this study is that participants were low-intermediate learners of French from vocational schools, which represents an underresearched population in the literature on the role of audio-visual input but also in SLA and CALL research in general.

2. Literature Review

2.1 Incidental Vocabulary Learning through Watching Videos with L1 and L2 Subtitles

Several studies have compared the effects of short exposure—that is between two and 60 minutes—to captioned or subtitled videos on L2 learners' incidental acquisition of vocabulary, as measured by tests that target meaning recognition, form recognition (Bianchi & Ciabattoni, 2008; Birulés-Muntané & Soto-Faraco, 2016; Vulchanova, Aurstad, Kvitnes, & Eshuis, 2015) and meaning recall (Peters et al., 2016).

Even though the majority of studies have reported learning gains at the level of immediate or delayed meaning recognition tests (i.e., recognize the L2 meaning among different L1 choices), they did not find differences between the effectiveness of captions and subtitles. Nonetheless, subtitles seemed slightly more beneficial in some studies (Bianchi & Ciabattoni, 2008; Peters et al., 2016; Vulchanova et al., 2015). In Birulés-Muntané and Soto-Faraco (2016), however, the captions group performed better than the subtitles group on an immediate meaning recognition test after watching an episode of a popular soap, but not significantly.

With regard to form recognition and meaning recall (i.e., recognize an L2 word form and provide an L1 translation respectively), the differential impact of captions and subtitles does not seem much stronger. In Peters et al. (2016), intermediate students watched a 13-minute documentary and received form recognition and meaning recall tests. The captions group scored significantly higher than the subtitles group on the form recognition test, whereas no difference could be established on the meaning recall test. Vulchanova et al. (2015), however, did not establish significant differences on their delayed form recognition test.

There are several gaps that need to be addressed. First, findings with regard to the effects of captions and subtitles on vocabulary learning are inconclusive. Second, learning gains in the studies mentioned previously were quite low (e.g., one to four words in Birulés-Muntané & Soto-Faraco, 2016 and Peters et al., 2016). This is undoubtedly due to the short time of exposure in the studies. Very few studies have looked into the potential of longer video treatments for L2 vocabulary learning. Some exceptions are Frumuselu, De Maeyer, Donche, and Gutiérrez-Colon Plana (2015), Pujadas and Muñoz (2019), and Zarei (2009).

In Zarei's (2009) study, three groups (captions, subtitles, and reversed) of L2 learners watched nine episodes of a popular series. They received meaning recognition and recall tests. The results demonstrated that, after the nine-week treatment, all groups recognized and recalled approximately 20 words, and that the captions group outperformed the two other groups significantly on

the meaning recall test. In Pujadas and Muñoz (2019), on the other hand, captions and subtitles groups did not differ on their meaning recall results after viewing 24 episodes of a series. Both groups did, however, recall up to 10% of the words they could learn.

In Frumuselu et al. (2015), university students (A2 to C1 proficiency levels) watched 13 captioned or subtitled episodes of *Friends* over a seven-week period. Participants received an immediate meaning recognition posttest including open and multiple-choice questions. Their reports of results did not enable to determine the number of words learned, but showed that the captions group scored higher than the subtitles group, independent of the learners' proficiency level.

Taken together, it seems that captions may be slightly more helpful for form recognition and meaning recognition, irrespective of the amount of exposure, whereas the results do not allow us to conclude anything for meaning recall yet. However, given the mixed results in previous research and the importance of exposure to audio-visual input for vocabulary learning (Peters, 2018), more research is needed on longer time of exposure to captioned and subtitled videos for L2 vocabulary learning.

2.2 Incidental Vocabulary Learning through Watching Videos: Learner- and Word-Related Factors

An increasing number of studies indicate that vocabulary learning from video is mediated by learner- and word-related factors (e.g., Peters, 2019; Peters & Webb, 2018). This highlights the need to take these aspects into account when investigating vocabulary learning through audio-visual input.

Prior Vocabulary Knowledge

Prior vocabulary knowledge refers to an estimate of the number of words of which learners have some knowledge. It is typically evaluated with a receptive test measuring learners' recognition of the meaning of a selection of words (see https://www.lextutor.ca/tests/ for some examples) and provides an estimate of learners' overall language proficiency.

Several studies examining the impact of watching captioned videos on learning gains demonstrated a large influence of prior vocabulary knowledge (Montero Perez et al., 2018; Peters et al., 2016; Peters & Webb, 2018). In Peters and Webb (2018), for example, every five additional known words on the prior vocabulary knowledge test would make the odds of recalling a target word three times higher. In Peters et al. (2016), knowing 10 additional words on the prior vocabulary knowledge test would increase the odds of learning a new word by 30%. Given the important role of vocabulary knowledge for word

learning from video and the lack of studies investigating its role with multiple videos (e.g., Rodgers, 2013), this is a crucial factor to consider.

Frequency of Occurrence

Frequency of occurrence of the target words in the input has been found to stimulate incidental vocabulary learning when reading (Pellicer-Sánchez, 2016) or listening (Vidal, 2011). However, little is known about its influence on incidental learning through videos. Rodgers (2013) reported a medium effect of frequency on vocabulary gains in a 10-week intervention study, whereas Peters and Webb (2018) as well as Peters et al. (2016) found a considerable effect of frequency on vocabulary acquisition with students who viewed only one short video. This seems to indicate that the role of frequency might depend on the number of videos watched. Yet, more research is needed to shed more light on this phenomenon.

Cognateness

Cognates, defined as words that are related in form and meaning (e.g., *motivation* [French] – *motivatie* [Dutch]), are assumed to facilitate learning gains (Peters, 2019). From the results of Vidal's (2011) study—which analyzed the impact of cognateness on listening and reading scores—it appeared that cognateness had a greater influence on the learning gains from listening than the gains from reading. Given that captioned/subtitled videos offer both modalities, it is important to consider their impact on vocabulary learning through audio-visual input, as was also indicated by Peters (2019) and Peters and Webb (2018).

Imagery

Imagery, or the visual representation of a word in a video, has the potential to influence learning gains (Rodgers, 2018). This is in line with Mayer's (2001) *Cognitive Theory of Multimedia Learning*, holding that words are more easily learned when simultaneously presented in images and words. We are only aware of one empirical study that has investigated the effects of imagery on vocabulary acquisition from video (Peters, 2019). In that study, intermediate L2 learners watched a 12-minute fragment twice and completed a form recognition and meaning recall test. A positive effect of imagery was found on both tests, supporting Mayer's (2001) theory. Yet, given the scarcity of research on this topic and the potentially important role of images for different aspects of vocabulary learning from video, more research is needed.

3. Rationale and Research Questions

In order to address the gaps identified in the literature review, this study focuses on the effects of viewing 135 minutes of authentic material over a four-week period. It investigates the influence of subtitles and captions and the role of learners' prior vocabulary knowledge as well as word-related factors on L2 incidental vocabulary learning. The specific research questions are:

1. How do captioned and subtitled videos impact incidental vocabulary learning of vocational school students after a four-week intervention program?
2. How is incidental vocabulary learning from video mediated by prior vocabulary knowledge, frequency of occurrence, cognateness, and imagery?

In this study, potential vocabulary learning gains are considered incidental because they are the results of a meaning-focused task where no attention is directed towards vocabulary (Webb, 2019).

4. Methodology

4.1 Participants

Vocational school students were recruited because they are not frequently represented in SLA/CALL research. One hundred and fifteen participants (114 males, 1 female; M_{age} = 17.47; *SD* = .64) from three Flemish vocational secondary schools took part in the study. The participants' L1 was Dutch. Data from five French-speaking students were removed from the dataset. They had two hours of formal French instruction per week and were considered to be low-intermediate learners of French based on their scores on a vocabulary knowledge test (see section 5.1).

Participants from two schools were randomly assigned to the captioned video (N= 39) or the subtitled video group (N= 36). Students from the third school represented the control group and only took the tests (N= 40). Since sessions took place over four weeks, data loss was inevitable. Twenty-four participants in the treatment groups were absent in at least one of the 11 sessions. Only the data from participants who attended all the sessions were used for analysis (N= 86).

4.2 Materials

Selection of Videos

We selected 15 videos from Belgian and French TV channels on the following topics: looking for a job, new technologies, and looking for or constructing a new house. An important criterion for the selection of the videos was that their themes had to correspond to the topics dealt with in class. In addition,

only authentic material designed for a French-speaking audience was included (Nunan, 2002).

Next, the videos were transcribed in French and translated into Dutch. Captioned and subtitled versions were created for all videos and were embedded in a website that was created for the experiment.

The vocabulary profile of each video was analyzed using *Lextutor* (www.lextutor.ca/vp/comp). Analyses demonstrated that 95% of the lexical coverage, which would be needed for global comprehension (Webb & Rodgers, 2009), was attained at 3000 most frequent words for four videos and at 4000 or 5000 most frequent words for 11 videos. While the vocabulary profile of the videos seems challenging for participants in the present study (i.e., low-intermediate learners), it has been argued that learners may achieve adequate comprehension of video with lower coverage percentages thanks to the presence of imagery (Rodgers, 2018).

Selection of Target Words

Target word selection was based on three criteria. First, participants completed a prior vocabulary knowledge test (i.e., VocabLab test, Noreillie, 2019; Peters, Velghe, & Van Rompaey, 2019), which measures learners' recognition of a selection of words pertaining to different frequency bands. The selected target words were in frequency bands that were not mastered by the students[1], namely the 1500–6000 word families. Second, the textbooks that they had been working with since the start of the school year were analyzed to ensure they did not contain the selected target words.

Third, since this study investigates the role of different word-related factors, words had to differ with regard to the parameters under investigation: frequency of occurrence, cognateness, and imagery. The target words appeared between 1 and 25 times in the videos. Cognates were evaluated by three Dutch-speaking yet highly proficient learners of French. A word was considered a cognate when at least two of the three raters labelled it as cognate (13 out of 50 target words). Finally, the visual representation of the target words was also evaluated. A word was considered as visually represented if the object/action it depicted appeared within 5 seconds before or after the audio and the on-screen text (Rodgers, 2018; 13 out of 50 target words). This resulted in a final set of 50 target words.

4.3 Instruments

Prior Vocabulary Knowledge Test

All groups received a 120-item multiple-choice prior vocabulary knowledge test two months before the treatment sessions. The test contains four different

frequency bands (0–4,000 most frequent words) with 30 items per frequency band. Each multiple-choice item contains the correct translation, three distractors, and an "I don't know" option. The internal reliability for the test was high (Cronbach's alpha = .91).

Content Comprehension Tests

Since vocabulary gains were meant to be incidental, a comprehension test was developed for each video, consisting of short open-ended, multiple-choice, and true/false questions. Questions were in participants' L1 in order not to interfere with L2 reading skills. Each test included two to five questions. The comprehension tasks were used to encourage students to watch the videos attentively and to determine whether they achieved adequate comprehension.

Vocabulary Tests

In order to test different aspects of vocabulary knowledge, participants completed two written vocabulary tests, which were administered as pretests (two weeks prior to the treatment) and at the end of the treatment (posttests).

Form recognition and meaning recall. The first test participants received measured their form recognition and meaning recall and consisted of the 50 target words and 30 distractors that were as frequent or more frequent than the target words. Participants had to indicate (by ticking off "yes" or "no") whether they had ever seen/heard the items (i.e., form recognition) and to give a translation or description of the word (i.e., meaning recall, see Figure 1). The reliability values of the form recognition pretest (Cronbach's alpha = .90) and posttest (Cronbach's alpha = .92), as well as of the meaning recall pretest (Cronbach's alpha = .88) and posttest (Cronbach's alpha = .87) were good.

Meaning recognition. The second vocabulary test participants received was a multiple-choice meaning recognition test (see Figure 1). For the same 80 words as in the first test (50 targets, 30 distractors), participants had to choose the correct translation from four options and an "I don't know" option. The

Figure 1. Example of vocabulary tests item.

different distractors in each MC item were in the same or a lower frequency band as the target word (Nation & Webb, 2011). The reliability of both meaning recognition tests was satisfactory (Cronbach's alpha = .87 for pretest and Cronbach's alpha = .89 for posttest).

Questionnaire

During the final session, the participants filled out two questionnaires. The first one focused on the participants' language background and on their contact with French outside the classroom. In the second questionnaire, participants had to rate the difficulty of the videos, the usefulness of the captions/subtitles, and had to indicate whether they had expected to receive the vocabulary posttests, in order to evaluate to what extent their learning was incidental.

4.4 Procedure

Two months before the experiment, the videos were pilot tested with 13 students in the same year and discipline as the actual participants of this study. Results indicated that participants found the videos interesting and comprehensible, but that captions or subtitles were helpful for comprehension.

At the same time, participants completed the prior vocabulary knowledge test. Two weeks before the experiment, they took the pretests that assessed learners' knowledge of the target words. They first received the form recognition and meaning recall and finally the meaning recognition test. This order was used so that the meaning recognition test could not help them for the meaning recall test. Each test took place during their regular French class and participants needed approximately 20 minutes to complete each test.

The video-based intervention lasted four weeks, with two sessions of approximately 30 minutes per week for the treatment groups (20 minutes of videos and 10 minutes of test). The experiment took place in a computer room with a PC and headset for each participant. Students had to watch the videos without taking notes. After watching the videos of one session, they completed the corresponding comprehension test. They were not allowed to watch the videos while answering the questions. During the final session, participants completed two unannounced vocabulary posttests in the same order as the pretests. After the vocabulary tests, they completed two questionnaires. They were then informed about the precise goal of the study.

Participants in the control group took part in three sessions: one for the prior vocabulary knowledge test, one for the pretest, and one for the posttests (four weeks later). They did not watch the videos. As specified by Nation and Webb (2011), a control group is needed when using real words as target words in order to control for learning gains outside of the treatment and to determine

whether the learning gains of the treatment groups can be ascribed to the videos and not to testing effects.

4.5 Scoring

Content Comprehension Tasks

One point was given for each correctly answered element of the open-ended questions. For the true–false and multiple-choice questions, one point was obtained for each correct answer.

Vocabulary Tests

In the form recognition pre- and posttests, all data were scored dichotomously with 0 for words ticked off as "not seen/heard" and 1 for words ticked off as "seen/heard". For the meaning recognition and recall tests, a word was considered "*known*", "*not known*" or "*learned*". A word was "*known*" if participants knew the word in the pretest and answered correctly in the posttest, as opposed to "*not known*" which corresponded to a correct answer in the pretest but incorrect in the posttest or incorrect in both pre- and posttests. A word was "*learned*" if the participant answered incorrectly in the pretest but correctly in the posttest (Horst, Cobb, & Meara, 1998). By making this distinction, we obtain absolute gains—i.e., learned words—and a relative gains percentage. Relative gains are calculated for each subject as follows:

$$\frac{\textbf{\textit{absolute gains}}}{\text{Number of target words} - \textbf{\textit{known}} \text{ words}} \times 100$$

Relative gains were used to account for the variation in participants' scores in the pretests. A participant who knew more words in the pretests had less room for improvement than a participant who knew fewer words in the pretests. Considering only the absolute gains might not highlight this difference.

4.6 Analyses

In answer to the first research question investigating the effects of captions and subtitles on vocabulary learning gains, one ANOVA per posttest was computed with relative gains as the dependent variable, and condition (control, captions, or subtitles) as the independent variable.

In answer to the second question on the role of learner- and word-related factors for vocabulary learning, two repeated measures logistic regression were performed in SPSS with the Generalized Estimating Equations procedure. This analysis is appropriate for dichotomous response data and allows

to account for the effect of word- and learner-related variables on participants' learning gains in one model. This analysis is based on the number of observations per participant. In other words, each observation is defined by a specific score (correct/incorrect) on a specific item for a specific participant. The odds of a correct response with a particular parameter are predicted by the odds ratio (i.e., $\exp^{b}$). This analysis thus provides the odds of a correct response for the following parameters: condition (captions or subtitles), prior vocabulary knowledge, frequency, cognateness and imagery. Since the treatment lasted four weeks, we controlled for a possible recency effect—i.e., whether the target items' last occurrence was in the last two weeks or not. The model was first run with all parameters, namely condition, prior vocabulary knowledge, recency, frequency, cognateness, and imagery. Next, nonsignificant parameters were removed and the model was refitted.

5. Results

5.1 Prior Vocabulary Knowledge

The descriptive statistics are displayed in Table 1. An ANOVA with condition as the independent variable revealed a significant difference between the groups ($F(2, 83) = 5.30$, $p = .007$, $\eta_p^2 = .11$). The post-hoc Bonferroni test indicated that the captions group significantly outperformed the control group ($p = .007$).

Table 1
Descriptive Statistics Prior Vocabulary Knowledge Test

Condition	Total score Mean % (SD)	0–1000 Mean % (SD)	1000–2000 Mean % (SD)	2000–3000 Mean % (SD)	3000–4000 Mean % (SD)
Captions (N = 28)	52.02% (13.55)	68.93% (3.26)	56.9% (3.89)	45.47% (4.01)	36.8% (4.34)
Subtitles (N = 22)	46.55% (11.4)	63.5% (3.54)	52.27% (3.91)	38.5% (2.87)	31.97% (4.70)
Control group (N = 36)	43.42% (12.32)	60.27% (3.61)	47.59% (4.08)	37.03% (3.69)	28.79% (4.33)

5.2 Content Comprehension Tests

In order to determine if participants understood the content of the videos, the mean result of all sessions was calculated. Both groups scored higher than 63%, with the subtitles group outperforming the captions group ($t(50) = -5.03$, $p < .001$, $\eta_p^2 = 0.33$). This was considered indicative of adequate comprehension

as determined by the A2 DELF official listening test for French as a second language threshold for success (Noreillie, 2019).

5.3 Questionnaire

Analysis of the questionnaire showed that 89% of the students found that the experiment was a pleasant activity and 60% evaluated the videos as easy to understand. When asked whether they would have understood the videos without captions/subtitles, 36% agreed while 63% did not (on a Likert-scale from 1= no opinion, 6= totally agree; $M = 3.22$, $SD = 1.13$). Importantly, only 10 students claimed to have understood the experiment was on vocabulary learning. However they did not score higher on the vocabulary tests.

5.4 RQ1: Differential Effects of Watching Videos with Captions or Subtitles on Participants' Vocabulary Learning Gains after a Four-week Intervention Program

Form Recognition

On the form recognition posttest, participants did not recognize more than half of the target words. More importantly, they falsely recognized half of the distractors that did not appear in the videos. We considered this a high-guessing rate and therefore excluded the results for further analyses.

Meaning Recall

For the meaning recall test, each group improved from pre- to posttest. As can be seen in Table 2, the treatment groups have greater absolute and relative gains than the control group.

Table 2
Descriptive Statistics Meaning Recall and Recognition Tests

	Meaning recall		**Meaning recognition**	
Conditions	**Absolute gains Mean (SD)**	**Relative gains (%) Mean (SD)**	**Absolute gains Mean (SD)**	**Relative gains (%) Mean (SD)**
Captions (N = 28)	5.00 (1.85)	14.02 (6.02)	5.57 (2.33)	21.81 (10.16)
Subtitles (N = 22)	4.59 (2.79)	12.16 (7.60)	5.82 (2.99)	20.77 (11.68)
Control group (N = 36)	3.25 (2.39)	9.07 (6.81)	3.75 (1.65)	13.57 (5.88)

An ANOVA[2] with condition as the independent and relative gains as the dependent variable yielded a significant effect of condition, $F(2, 83) = 4.36$, $p = .016$, $\eta_p^2 = .35$. The post-hoc Bonferroni tests revealed that only the captions group significantly outperformed the control group ($p = .014$).

Meaning Recognition

Similar to the meaning recall test, an ANOVA with condition as the independent variable and relative gains as the dependent variable was run (see Table 2 for descriptive statistics). Since there was no homogeneity of variance ($p = .001$), the obtained Welch's adjusted F ratio (9.34), reported as $F(2, 41) = 7.76$, $p < .000$, $\eta_p^2 = .158$, was used.

Games Howell post-hoc results indicated a significant difference between the captions and the control groups ($p = .001$) and between the subtitles and control groups ($p = .031$). The two treatment groups' scores did not significantly differ.

5.5 RQ2: Relationship Between Vocabulary Learning Through Watching Videos and the Learner- and Word-Related Variables

Meaning Recall

Since the subtitles group did not significantly outperform the control group on the meaning recall test, only the data of the captions group were taken into account. A GEE analysis was conducted for 1019 observations (see Table 3), i.e., items unknown in the pretest which could be learned in the treatment. As can be seen in Table 4, frequency, imagery, cognateness, and prior vocabulary knowledge had a significant impact on vocabulary learning through watching video.

Regarding the role of frequency of occurrence, the analysis showed that, for each new encounter with a target item in the videos, the odds of a correct answer increased by 7%. Imagery increased the odds of recalling the item by 1.5 and the odds were 3 times higher if the word was a cognate. Finally, one additional correct answer on the prior vocabulary knowledge test increased the odds of a correct answer in the posttest by 1%.

Table 3
Number of Incorrect/Correct Responses in the Meaning Recall Test

	Correct responses	Incorrect responses	Total
Captions group	140 13.7%	879 86.3%	1019 100%

Table 4
GEE: Meaning Recall Test

						CI	
Parameter	**Wald-Chi Square**	**df**	**Sig**	**B**	**Exp(B)**	**Lower**	**Upper**
Intercept	160.048	1	.000	-3.610	.027	.015	.047
Frequency of occurrence	36.777	1	.000	.071	1.074	1.049	1.098
Imagery	5.946	1	.015	.457	1.580	1.094	2.282
Cognateness	27.682	1	.000	1.286	3.618	2.241	5.841
Prior vocabulary knowledge	13.617	1	.010	.014	1.014	1.007	1.022

Meaning Recognition

The GEE analysis for meaning recognition was based on both experimental groups data resulting in 1400 observations (see Table 5). Prior vocabulary knowledge, frequency, and cognateness correlated positively with participants' learning gains (see Table 6).

Table 5
Number of Incorrect/Correct Responses in the Meaning Recognition Test

	Correct responses	**Incorrect responses**	**Total**
Treatment groups	284	1116	1400
	20.3%	79.7%	100%

Table 6
GEE: Meaning Recognition Test

						CI	
Parameter	**Wald-Chi Square**	**df**	**Sig**	**B**	**Exp(B)**	**Lower**	**Upper**
Intercept	33.979	1	.000	-2.617	.073	.030	.176
Prior vocabulary knowledge	5.360	1	.021	.016	1.016	1.002	1.029
Frequency of occurrence	17.190	1	.000	.043	1.044	1.023	1.066
Cognateness	6.430	1	.011	.451	1.570	1.108	2.225

The analysis indicated that for each additional known word in the prior vocabulary knowledge test, the odds of a correct response on the meaning recognition test increased by 2%. The odds of a correct response on the meaning recognition test increased by 4% for each additional occurrence of the target word. Learners' odds of a correct answer were 1.5 times higher when the target was a cognate.

6. Discussion

This study measured the impact of a four-week intervention program on the incidental vocabulary learning gains of low-intermediate learners of French from vocational schools. In addition, it investigated the role of captions and subtitles as well as learner- and word-related factors, namely prior vocabulary knowledge, frequency, imagery, and cognateness.

6.1 Effects of Watching Captioned and Subtitled Videos on Vocabulary Gains after a Four-week Treatment

Findings demonstrated that exposure to subtitled or captioned video over a four-week period led to on average 15% of the target words learned. It also showed that subtitles led to better comprehension than captions, which is in line with other studies on comprehension through watching videos (e.g., Bianchi & Ciabattoni, 2008). However, our results did not allow us to define a clear-cut differential effect of captions and subtitles on the aspects of word knowledge measured, since we found no significant difference between the treatment groups on both meaning recall and recognition tests.

There are, however, indications that captions were more beneficial for the meaning recall test, as also found in Zarei (2009), since only the captions group outperformed the control group on this test. This difference might result from the written test format. Indeed, participants of the captions group could not only hear the French words but also see the correct written forms of the words in the captions, as opposed to the subtitles group. Another plausible explanation is that students from the subtitles group focused mainly on the Dutch subtitles when discourse was hard to understand and consequently paid less attention to the French discourse, which hindered the initial form–meaning mapping. This hypothesis is supported by the students' answers on the questionnaire, which indicated that the discourse was sometimes too fast to follow. Finally, the imagery may have favored the captions group more than subtitles group. Indeed, the captions group could rely on the aural, written, and visual representation of the words to establish the form–meaning link. When confronted with the visual clues, participants in the subtitles group

might have encountered difficulties linking the L1 written form to the L2 aural and visual representations.

While learning gains might seem low, they are in line with other findings in incidental learning studies. Moreover, it is possible that participants learned other words from the videos that were not included in the posttest. In comparison, in Rodgers (2013) participants recognized on average six words after 420 minutes of exposure to a TV program. While Vanderplank (2010) argued that captions might be more profitable for high-proficient learners, our results demonstrated that even low-intermediate learners from vocational schools can benefit from captioned videos.

6.2 Role of Learner- and Word-related Factors

Results showed that the role of learner- and word-related factors differed in function of the vocabulary aspect measured. While prior vocabulary knowledge, frequency of occurrence, and cognateness influenced both meaning recognition and recall, imagery only affected meaning recall.

Prior Vocabulary Knowledge

The influence of prior vocabulary knowledge on vocabulary gains was minimal. While our findings reveal a positive relationship, they are less pronounced than in Peters et al. (2016), since they found that, for each additional correct answer on the prior vocabulary knowledge test, the odds for a correct answer on the meaning recall test increased by 3%, compared to 1.4% in our study. Other studies that took vocabulary size into account in their analysis (Montero Perez et al., 2018; Peters & Webb, 2018) also found a greater positive influence on the vocabulary gains. The limited influence of prior vocabulary knowledge could be due to the small variation in scores on the prior vocabulary knowledge test which might make it difficult to see whether and how vocabulary knowledge mediates learning.

Frequency

Findings of the present study showed that the impact of frequency was limited. This contrasts with results of the studies of Peters et al. (2016) and Peters and Webb (2018), where frequency played an important role. However, participants of these two studies watched only one video and were tested right after the treatment. This means that the range of occurrence, that is the number of videos the word appeared in, was smaller in their study. In the present study, almost all the target items appeared in at least two different videos across four weeks. Therefore, it is possible that the effect of frequency was diminished by the time between the appearances (e.g., a word could occur three times in one

video and three times in another one two weeks later). However, it should also be mentioned that no recency effect was found, meaning that words that were most learned were not only words that occurred in the last two weeks of the intervention.

Cognateness

Cognateness had the strongest influence on the learning gains found in this study, with 3 times more odds to recall the meaning and 1.5 times more odds to recognize the meaning of a cognate than non-cognate. Nonetheless, even if the impact is substantial, it is still a smaller effect compared to the results of Peters and Webb (2018). They observed that the odds to recall and recognize a cognate were eight times and 2.5 times higher respectively. However, as they suggest themselves, this could be a consequence of the relatedness between the languages used in their study, English and Dutch. As Dutch is less related to French than to English, this could have had an effect on the role of cognateness. While the present results support the importance of cognateness in learning vocabulary through videos, further research on this variable is strongly encouraged.

Imagery

Another factor that facilitated vocabulary learning is imagery, since it moderated the learning gains in the meaning recall test, as in Peters (2019), which indicates that the visual clues helped learners establish form–meaning links in the mental lexicon. These findings corroborate Mayer's (2001) *Cognitive Theory of Multimedia Learning* and are also in line with Bianchi and Ciabattoni (2008), which reveals that beginner learners in the captions group benefited from the combination of captions and imagery. However, the reason why it influenced the results of the meaning recall but not of the meaning recognition test is puzzling. It could be argued that the difficulty of the test plays a role in this. A meaning recall test is more demanding than a meaning recognition test. Therefore, imagery might only have played a central role in a later stage of the form–meaning link in the mental lexicon. Participants did not need the imagery to be able to recognize the word, however, when needing to recall it, remembering the visual clue might have favored the process. Yet, more research is warranted in order to further clarify the role of imagery, as also highlighted by Peters (2019) and Rodgers (2018).

7. Limitations

This study has some limitations. First, to make sure that cognates were really considered cognates by our learners, it could have been better to let learners

of the same proficiency evaluate them. However, they have been evaluated in the same way as in Peters and Webb (2018).

Second, some participants reported having understood the main goal of the experiment, which is incidental vocabulary learning. However, we still considered any learning gains as incidental and not intentional, since participants were never informed about the vocabulary posttests and the treatment was a meaning-focused task where no attention was directed towards vocabulary.

In addition, considering the important amount of guessing in the form recognition test, it may be hypothesized that it was not adapted to the study design—i.e., longer treatment than in other studies using this test (e.g., Peters et al., 2016). Because the test did not include an "I don't know" option nor any pseudowords, the amount of guessing could not be decreased.

Finally, it could be argued that the videos were too difficult for the participants. However, the comprehension tests showed that adequate comprehension was attained. Moreover, even though learners did not reach 95% coverage, they were still able to learn new words while viewing. Yet, videos with a lower coverage might have led to more learning gains.

8. Conclusion

This study expands on existing research into the role of audio-visual materials for L2 learning by (1) researching the role of captioned and subtitled videos on vocabulary learning during a four-week period and (2) including different factors that influence vocabulary learning through watching videos: prior vocabulary knowledge, frequency, cognateness, and imagery. Results reveal that participants picked up new words while watching videos and that frequency of occurrence, cognateness, and prior vocabulary knowledge mediated learning gains on both meaning recall and recognition tests, while imagery influenced the meaning recall test results only.

Several pedagogical implications emerge from this intervention. This study investigated the learning gains of low-intermediate students of vocational schools and therefore demonstrates that watching videos is not only a fruitful L2 learning activity for high-proficient learners. Moreover, learners usually appreciate this activity, as also reported by our participants. This should encourage teachers to include videos both inside the classroom and as out-of-class activities.

Finally, teachers should be encouraged to use subtitled videos if they want to stimulate initial form–meaning connections—meaning recognition—and comprehension. However, they might want to use captioned videos to strengthen this form–meaning link. Importantly, they should also consider

different factors such as frequency of occurrence, cognateness, and imagery when selecting materials in order to maximize incidental vocabulary learning gains.

Notes

1. Participants needed a score of 27 out of 30 to master a level in the test (Noreillie, 2019).
2. Homogeneity of variance was assumed, $p = .289$

About the Authors

In her doctoral research, Isabeau Fievez investigates the effectiveness of videos for different aspects of vocabulary knowledge. She is doing her PhD as part of an INTERREG V project.

Maribel Montero Perez is Assistant Professor at Ghent University, Belgium. Her research interests include, amongst others, multimodal input for L2 learning, support tools for L2 learners in online learning environments, technology-enhanced listening, and vocabulary learning.

Frederik Cornillie (PhD) is research and valorization manager at ITEC. His main research interest concerns the effectiveness and use of learning support in technology-mediated environments, in particular in relation to individual differences.

Piet Desmet is a full professor of French, applied linguistics and CALL at KU Leuven, campus Kulak Kortrijk, and coordinator of ITEC, imec research group at KU Leuven.

References

Bianchi, F., & Ciabattoni, T. (2008). Captions and subtitles in EFL learning: An investigative study in a comprehensive computer environment. In A. Baldry, M. Pavesi, & C. Taylor Torsello (Eds.), *From Didactas to Ecolingua* (pp. 69–80). Trieste, Italy: Edizioni Università di Trieste.

Birulés-Muntané, J., & Soto-Faraco, S. (2016). Watching subtitled films can help learning foreign languages. *PLoS ONE*, *11*(6), 1–10. https://doi.org/10.1371/journal.pone.0158409

Frumuselu, A. D., De Maeyer, S., Donche, V., & Gutiérrez-Colon Plana, M. (2015). Television series inside the EFL classroom: Bridging the gap between teaching and learning informal language through subtitles. *Linguistics and Education*, *32*, 107–117. https://doi.org/10.1016/j.linged.2015.10.001

Horst, M., Cobb, T., & Meara, P. (1998). Beyond a Clockwork Orange: Acquiring second language vocabulary through reading. *Reading in a Foreign Language*, *11*(2), 207–223.
Mayer, R. E. (2001). *Multimedia Learning* (2nd ed.). New York, NY: Cambridge University Press. https://doi.org/10.1017/CBO9781139164603
Montero Perez, M., Peters, E., & Desmet, P. (2018). Vocabulary learning through viewing video: The effect of two enhancement techniques. *Computer Assisted Language Learning*, *31*(1–2), 1–26. https://doi.org/10.1080/09588221.2017.1375960
Montero Perez, M., & Rodgers, M. (2019). Video and language learning. *The Language Learning Journal*, *47*(4), 403–406. https://doi.org/10.1080/09571736.2019.1629099
Nation, P. (2013). *Learning vocabulary in another language* (2nd ed.). Cambridge, England: Cambridge University Press.
Nation, P., & Webb, S. (2011). *Researching and analysing vocabulary*. Boston, MA: Heinle.
Noreillie, A. (2019). *It's all about words: Three empirical studies into the role of lexical knowledge and use in French listening and speaking tasks*. (Unpublished dissertation) Catholic University of Leuven, Belgium.
Nunan, D. (2002). Listening in language learning. In J. C. Richards & W. A. Renandya (Eds.), *Methodology in language teaching* (pp. 238–241). Cambridge, England: Cambridge University Press.
Pellicer-Sánchez, A. (2016). Incidental L2 vocabulary acquisition from and while reading: An eye-tracking study. *Studies in Second Language Acquisition*, *38*(1), 97–130. https://doi.org/10.1017/S0272263115000224.Incidental
Peters, E. (2018). The effect of out-of-class exposure to English language media on learners' vocabulary knowledge. *ITL - International Journal of Applied Linguistics*, *169*(1), 142–168. https://doi.org/10.1075/itl.00010.pet
Peters, E. (2019). The effect of imagery and on-screen text on foreign language vocabulary learning from audiovisual input. *TESOL Quarterly*, *53*(4), 1008–1032. https://doi.org/10.1002/tesq.531
Peters, E., Heynen, E., & Puimège, E. (2016). Learning vocabulary through audiovisual input: The differential effect of L1 subtitles and captions. *System*, *63*, 134–148. https://doi.org/10.1016/j.system.2016.10.002
Peters, E., Velghe, T., & Van Rompaey, T. (2019). The VocabLab tests: The development of an English and French vocabulary test. *ITL - International Journal of Applied Linguistics*, *170*(1), 53–78. https://doi.org/10.1075/itl.17029.pet
Peters, E., & Webb, S. (2018). Incidental vocabulary acquisition through viewing L2 television and factors that affect learning. *Studies in Second Language Acquisition*, 1–27. https://doi.org/10.1017/S0272263117000407
Pujadas, G., & Muñoz, C. (2019). Extensive viewing of captioned and subtitled TV series: A study of L2 vocabulary learning by adolescents. *The Language Learning Journal*, 1–18. https://doi.org/10.1080/09571736.2019.1616806
Rodgers, M. (2013). *English language learning through viewing television: An investigation of comprehension, incidental vocabulary acquisition, lexical coverage, attitudes, and captions*. (Unpublished dissertation). Victoria University of Wellington.
Rodgers, M. (2018). The images in television programs and the potential for learning unknown words. *ITL - International Journal of Applied Linguistics*, *169*(1), 191–211. https://doi.org/10.1075/itl.00012.rod
Vanderplank, R. (2010). Déjà vu? A decade of research on language laboratories, television and video in language learning. *Language Teaching*, *43*(1), 1–37. https://doi.org/10.1017/S0261444809990267

Vidal, K. (2011). A comparison of the effects of reading and listening on incidental vocabulary acquisition. *Language Learning*, *61*(1), 219–258. https://doi.org/10.1111/j.1467-9922.2010.00593.x

Vulchanova, M., Aurstad, L. M. G., Kvitnes, I. E. N., & Eshuis, H. (2015). As naturalistic as it gets: Subtitles in the English classroom in Norway. *Frontiers in Psychology*, *5*, 1–10. https://doi.org/10.3389/fpsyg.2014.01510

Webb, S. (2019). *The Routledge handbook of vocabulary studies*. Abingdon, England: Routledge. https://doi.org/10.4324/9780429291586

Webb, S., & Rodgers, M. (2009). Vocabulary demands of television programs. *Language Learning*, *59*(2), 335–366. https://doi.org/10.1111/j.1467-9922.2009.00509.x

Zarei, A. A. 8j-level-three-STACKED8j-level-three-STACKED8j-level-three-STACKED8j-level-three-STACKED(2009). The effect of bimodal, standard, and reversed subtitling on L2 vocabulary recognition and recall. *Pazhuhesh-E Zabanha-Ye Khareji*, *49* (Special Issue), 65–85.