The effects of video caption types and advance organizers on incidental L2 collocation learning

(Mark) Feng Tenga,b,∗

a School of International Studies, Sun Yat-sen University, Zhuhai, China
b Department of Education Studies, Hong Kong Baptist University, Hong Kong, China

ARTICLE INFO

Keywords:
Full captions
Keyword captions
Advance-organizer strategies
Collocation
Incidental learning

ABSTRACT

This study examined the effects of captioned videos and an advance-organizer strategy in incidental learning of collocations. Videos were enhanced by (1) inserting various types of captions (i.e., full captioning, keyword captioning, and no captioning); (2) administering or not administering an advance-organizer activity prior to video viewing; and (3) the combined effects of these two variables. This study adopted a 2 (+/− advance-organizer strategy) × 3 (types of captioning) between-subjects design, resulting in six possible conditions. A total of 361 Chinese primary school students learning English as a second language (L2) each watched a series of four videos in one of the six conditions. Data from the post-intervention test revealed that full captioning led to better performance in incidental learning of collocations than keyword captioning or no captioning. Administering the advance-organizer strategy resulted in better collocation learning gains compared to no administration. Learners who completed the advance-organizer strategy in the full-captioning condition achieved the best results on incidental learning of collocations. Incidental learning of collocations occurs at a rate similar to the learning of single words. This article highlights the importance of fully captioned videos and an advance-organizer strategy for incidental learning of collocations among primary school L2 learners.

1. Introduction

Collocation—a type of multi-word unit—has been suggested as a prerequisite for achieving advanced levels of language proficiency and fluency in second language (L2) learning (Schmitt, 2010). Despite efforts to explore best practices in the teaching and learning of individual words, research on L2 vocabulary development has only recently started to examine learning performance with multi-word units, including collocations (Webb & Nation, 2017). L2 learners’ acquisition of multiple-word units have been found to lag behind single-word acquisition (e.g., Pellicer-Sánchez, 2017); even when learners can produce multi-word units, they may demonstrate inappropriate usage (e.g., Nesselhauf, 2005). Moreover, L2 learners may not have sufficient intuition to determine the frequency of collocations and then process or learn them (Siyanova-Chanturia & Spina, 2015). Schmitt (2010) contended that, compared to the acquisition of single words, collocations and other multi-word units pose unique challenges for L2 learners.

Various approaches including reading, listening, and read-while-listening have been employed to assess vocabulary learning. Recent advances in technology have rendered authentic multimedia language-learning materials more accessible to L2 learners, and captioned videos have been employed to enhance vocabulary learning (Montero Perez, Peters, & Desmet, 2018; Teng, 2019a). Captions, a form of on-screen text displayed alongside a same-language soundtrack (Danan, 2004), have been used increasingly to...
facilitate comprehension (Teng, 2019b) and encourage vocabulary learning (Peters, Heynen, & Puimège, 2016; Peters & Webb, 2018). Captioning may serve as an added bonus rather than a deducted merit for L2 learners, who may connect auditory input to visual input when learning novel words (Vanderplank, 2016).

Despite the potential of using captioned videos to learn individual L2 vocabulary words, learning L2 collocations may differ. For example, compared to students learning individual L2 words under a reading-while-listening condition (Teng, 2018a), learners gained limited collocation learning outcomes in the same condition (Webb, Newton, & Chang, 2013). The learning burden of a word is affected by the learner’s knowledge of the wider world and by features that render certain words and multi-word combinations memorable (Webb & Nation, 2017), suggesting difficulties in learning collocations (Nesselhauf, 2005). Learners may encounter challenges in obtaining accurate English collocations because they may be unable to discern how the meanings of separate parts relate to the meaning of the whole-word sequence (Chan & Liou, 2005). Whether L2 learners can notice and discern form-and-meaning links for new collocations from captioned video input is still unclear.

Limited classroom time makes the teaching of all collocations via explicit vocabulary activities impossible. Thus, it is essential to identify effective ways to maximize incidental learning of collocations, i.e., an approach of learning vocabulary by focusing learners’ attention on discerning the meanings of messages without deliberately committing the items to memory (Hulstijn, 2013); however, the incidental learning process has been shown to be incremental, for which L2 learners only gained limited outcome (Webb et al., 2013). Incidental learning of collocations through audiovisual inputs may be challenging because of rapid online processing demands while inferring meanings of collocations from literal concepts. Identifying productive ways to maximize the effects of captioned videos on learning collocations warrants consideration. To address the limited effects of captioned videos, an advance-organizer strategy has been suggested as an approach for strengthening L2 learners’ capacity to associate prior knowledge with audiovisual input in a multimedia environment for a better interpretation and comprehension of the material (Li, 2014). To the best of the author’s knowledge, incidental learning of collocations through the interplay of captions and advance-organizer strategies has not yet been explored.

In investigating incidental learning of collocations, the present study included two independent variables: either the presence or absence of an advance-organizer strategy, and the type of captioning in videos (i.e., full captioning, keyword captioning, or no captions). This study is novel in its collective evaluation of an advance-organizer strategy and captions. Current trends in technology-enhanced learning have highlighted adaptive learning (i.e., delivery of instruction addressing an individual’s unique needs through timely feedback, pathways, and resources) as an innovative technique compared to the traditional ‘one-size-fits-all’ learning experience. Given diversity in learning styles, cognitive styles, and preferred multimedia, it is important to understand adaptive learning and its role in shaping suitable learning materials. Findings provide insights for pedagogical decisions and enable L2 teachers to identify more effective methods for incidental learning of collocations in a multimedia classroom setting.

2. Literature review

2.1. Knowledge of L2 collocations

A collocation is defined as a combination of certain words, e.g., the habitual juxtaposition of a particular word with another word (Webb & Nation, 2017). Most combinations are not arbitrary but follow grammatical rules and the semantic requirements of the words involved (Liu, 2010). Collocations can be identified based on either semantic principles (Cowie, 1994) or the strength of co-occurrence of two lexical items (Sinclair, 1991). Drawing from Sinclair’s definition, Hoey (1991) described collocation as “the relationship a lexical item has with items that appear with greater than random probability in its (textual) context” (p. 7). Research has shown that collocations are particularly problematic for L2 learners, as these learners may struggle to determine the degree of convention underlying certain combinations (Pellicer-Sánchez, 2017).

2.2. Incidental vocabulary learning from captions

Captions refer to a stream of written text presented synchronously with video and audio reinforcement (Danan, 2004). The cognitive process involved in watching captioned videos has been described as bi-modal input. Such input can provide learners with multiple representations of the same information (Vanderplank, 2016). Theoretical frameworks supporting captions include Paivio’s dual-coding theory (1986) and Mayer’s multimedia principle (2001). In the former framework, verbal and imagery systems can work together with, act upon, store, and retrieve incoming information for subsequent use, thereby helping learners retrieve a word and visual clue simultaneously. In the latter framework, text and graphics may be more conducive to learning than either medium alone because learners’ understanding may be facilitated by cognitive processing while attending to material, deconstructing the material into a cognitive representation, and analyzing the material based on existing knowledge. These two frameworks highlight the integration of text and graphics and theoretically support the potential for captions to enhance L2 English learners’ vocabulary learning.

Empirical studies have been conducted to explore incidental vocabulary learning when watching captioned videos. Monterro Perez, Peters, Clarebout, and Desmet (2014) divided 133 Flemish undergraduate students into four captioning conditions: a no-captioning condition (n = 32); keyword-captioning condition (n = 34); full-captioning condition (n = 30), and full-captioning-plus-highlighted-keywords condition (n = 37). Findings revealed that keyword captioning and full captioning plus highlighted keywords condition were significantly more effective than full captioning and no captioning in terms of form recognition; however, comprehension and meaning recall were challenging for L2 learners across the four conditions. These results support the potential of captions in enhancing recognition of word form and meaning in addition to highlighting challenges in recalling word form and meaning.
(Neuman & Koskinen, 1992).

In a recent study (Montero Perez et al., 2018), learners in the keyword captioning and glossed keywords condition significantly outperformed those in the conditions of full captioning and no captions. However, Montero Perez, Peters, and Desmet (2013) revealed that learners in the full captioning condition outperformed those in the other two conditions (keyword captioning and no captioning) in terms of L2 listening. Teng (2019a) demonstrated that full captioning led to more pronounced incidental vocabulary learning compared to keyword captioning. Despite controversy regarding full captioning, this approach has demonstrated promise in enhancing incidental vocabulary learning. As explained by Winke, Gass, and Sydorenko (2010), the differential effects of captioning may be due to learners’ attention. Learners who attend actively to L2 forms often learn words better; however, when linguistic input from a captioned video is complex, a string of incomprehensible input may lead L2 learners to focus on particular parts of language rather than overall comprehension.

Other studies have challenged the use of captioned videos by showing that students were too reliant on reading captions and most ignored the soundtrack, resulting in only slight improvement in comprehension of the video content (Latifi, Mobalegh, & Mohammadi, 2011). Similarly, Diao, Chandler, and Sweller (2007) found that while watching captioned videos, L2 learners depended too much on reading on-screen text; they became passive to incoming information, which led to superficial learning of difficult words. Taylor (2005) argued that simultaneous presentation of pictures (animations) and words (on-screen text or captions) may cause visual input to overwhelm L2 learners due to a limited capacity to process audiovisual material.

The above findings imply the following: (a) the effects of captioned videos on L2 learning deserve further exploration, even though controversy remains around the use of captioned videos; (b) different types of captioning may influence incidental vocabulary learning differently; (c) learners may not have sufficient time to infer word meaning while viewing videos, rendering recall of word form and meaning more challenging than form and meaning recognition; and (d) incidental vocabulary learning has proven challenging, and only partial learning gains have been achieved. Given these issues, further investigation should be conducted to determine the potential roles of captions in incidental vocabulary learning.

No empirical evidence is available on using captioned videos for incidental learning of L2 collocations, but studies on incidental vocabulary learning using captioned videos can provide relevant insights. Captioned videos may be a powerful means of encouraging L2 students’ collocation learning. For example, captioned videos may help L2 learners develop more multidimensional and extensive understandings of new words and their collocations. This benefit, along with the possibility to orient L2 learners’ attention, may decrease cognitive demands and help learners seek textual meaning, leading to possible L2 collocation acquisition. The captioning types in the present study were similar to the three techniques of full captioning, keyword captioning, and no captions in Teng (2019a) and Montero Perez et al. (2013). The difference lay in the evaluation of collocation learning, which has been reported as more challenging than word learning (Pellicer-Sánchez, 2017). In addition, the focus was on young learners, which, based on Teng (2019a, 2019b), deserves more attention in the area of computers and education.

2.3. L2 learning from advance organizers

An advance organizer is a teaching tool used to orient a student to a new topic, provide an overview of new information, and illustrate the structure or organization of upcoming material (Li, 2014). Vandergrift (2007) suggested that using advance organizers helped learners develop a conceptual framework prior to learning through organizing and interpreting incoming new information. A systematic review of empirical studies on using advance organizers suggests various options for their design, including a summary of video plots with accompanying pictures, presentation of key vocabulary, pictorial illustrations, verbal descriptions, cultural background cues, and pre-viewing questions (Chung & Huang, 1998; Herron, 1994; Li, 2014). For instance, Chung and Huang (1998) suggested that Taiwanese English learners exposed to advance organizers prior to viewing videos demonstrated better video comprehension. In an earlier study (Herron, 1994), advance organizers were found to be effective in aiding video comprehension due to their offering learners a preview of the structures, vocabulary, and cultural information associated with the videos. Similarly, Li (2014) employed an interactive advance-organizer activity before watching a DVD video to afford L2 learners access to pictures, words, and spoken words in the videos. Based on a sample size of 97 learners in Taiwan, Li concluded that the group who watched captioned videos and used advance organizers outperformed other groups in listening comprehension (i.e., those who watched captioned videos only, used advance organizers only, and were in the control group).

Thus, advance-organizer strategies may be effective in activating learners’ prior knowledge (e.g., linguistic and content knowledge) and integrating new knowledge with prior knowledge to better interpret audio and visual input embedded in a multimedia technology setting (e.g., watching captioned videos). Relevant video information provided by advance organizers may help L2 learners engage in cognitive processing to comprehend aural input.

Research has identified the positive role of advance-organizer strategies in L2 listening comprehension (e.g., Li, 2014), although a limited number of studies have considered advance-organizer strategies in enhancing vocabulary knowledge. Ponce, Mayer, Figueroac, and López (2018) tested an interactive advance-organizer method through two experimental studies involving college students and high school students. Both studies included experimental groups using the advance-organizer method, wherein students were given a formative assessment in vocabulary knowledge after highlighting unknown words in a passage displayed on their computer screen compared to a control group using traditional treatment. Results suggested the roles of advance-organizers in enhancing vocabulary knowledge. However, it appears that few studies have measured the effects of advance-organizer strategies on incidental vocabulary learning, let alone incidental learning of collocations, an important but often neglected aspect of vocabulary acquisition (Webb & Nation, 2017). Collocations have become increasingly important as the improved availability of multimedia and digital technologies has rendered advance organizers more common, informative, and meaningful for L2 learners. This supposition
elicits future research to shape pedagogical decisions and help L2 teachers identify an alternative approach for effective instruction of learning collocations. Different from previous studies (e.g., Li, 2014; Ponce et al., 2018), the advance-organizer technique in this study included screenshots from videos and was designed as an online test system. Young learners may be more interested in this technique than traditional methods used in prior studies. Through this approach, learners may be better able to integrate prior knowledge and organize received information for improved comprehension.

2.4. The present study

The aforementioned studies revealed that using captions and advance organizers may enhance L2 learning. In contrast to such work, the present study contributes to this body of research by exploring the interplay between captions and an advance-organizer strategy. The aim was to discover the respective effects of L2 captions and advance-organizers as well as an optimal combination of the two variables on incidental learning of collocations. It is essential to use captioned videos and advance organizers to measure incidental learning, which serves as an important component of learners’ social and cognitive development. Compared to incidental learning, intentional learning requires that students learn material and commit it to memory. However, classroom learning is often limited to intentional learning. Yet there are other ways to acquire knowledge independent of conscious attempts to learn, e.g., the use of captions and advance organizers. To the author’s limited knowledge, no previous studies have focused on measuring incidental learning of collocations using captions or advance organizers. The results of the present study provide substantial insight into incidental learning of collocations, a relatively under-research dimension in vocabulary research. This study seeks to address the following research questions:

1. What is the main effect of an advance-organizer strategy on L2 young learners’ incidental learning of collocations?
2. What is the main effect of captioned videos on L2 young learners’ incidental learning of collocations?
3. What is the interaction effect between captioned videos and an advance-organizer strategy on L2 young learners’ incidental learning of collocations?

3. Method

3.1. Research design

The present study adopted a 2 × 3 between-subjects design. The two independent variables were captioning types (full captioning, keyword captioning, and videos without captions) (Montero Perez et al., 2013) and an advance-organizer strategy (Li, 2014). The two independent variables resulted in six experimental groups. Group 1 viewed full-captioning videos and completed advance-organizer exercises (FC + AO, n = 59). Group 2 viewed full-captioning videos without advance-organizer exercises (FC, n = 67). Group 3 viewed keyword-captioning videos and completed advance-organizer exercises (KC + AO, n = 54). Group 4 viewed keyword-captioning videos without advance-organizer exercises (KC, n = 61). Group 5 viewed videos without captions and completed advance-organizer exercises (NC + AO, n = 58). Finally, Group 6 viewed videos without captions and did not participate in advance-organizer exercises (NC, n = 62). Dependent variables included a post-test measuring four dimensions of collocation knowledge (see Measures section).

3.2. Participants

Six English teachers from six Direct Subsidy Scheme schools in Hong Kong assisted with this study. All teachers held a Postgraduate Diploma in Education (PGDE) in English education and had taught English for at least 5 years. PGDE is a program that prepares university graduates to teach in Hong Kong in sectors ranging from kindergartens to secondary schools. English was used as the main language of instruction in the schools except in Chinese courses. The six teachers co-designed an internal English test to measure students’ core English skills, including reading, writing, listening, and speaking. The internal test was administered to 100 sixth-grade students in each school prior to the study. 61, 57, 58, 68, 61, and 62 learners in each school earned 80–85 out of 100 points, and this large number of participants were selected. A few learners who were under 70 points or above 90 points were not selected. Analysis of variance (ANOVA) comparing the six groups’ performance on this test did not reveal a significant difference in scores \( F(5, 361) = 1.134, p = 2.89 \). Cronbach’s alpha (0.79) indicated acceptable reliability for the test. This internal test was modelled on Cambridge English: B1 Preliminary for Schools; benchmarked with the Common European Framework of Reference for Languages (CEFR) scale, participants’ English could be described as at the B1 level. Students at this level can understand the main points of input on familiar matters regularly encountered in work, school, leisure, etc. However, data from six students who did not follow the research procedure were not included in the final analysis. The final number of participants was 361 (59, 54, 58, 67, 61, and 62 in each school). The six schools were randomly assigned to one of six groups with one teacher per group. There were 185 girls and 176 boys (\( M_{age} = 11.82, SD = 2.11 \)), all of whom had received formal English instruction for at least 6 years. Their native language was Chinese (Cantonese or Putonghua), and English was their L2.
3.3. Materials

3.3.1. Videos

Four authentic episodes were selected from a series of YouTube storytelling videos for children. Each episode was approximately 10 min long, which allowed several collocations to be inserted. The difficulty in locating a sufficient number of unknown collocations in video texts for participants added a challenge to the study’s purpose of measuring incidental learning of collocations from videos. Hence, the text was modified to suit the research aims. Unknown collocations were inserted in the text. At least 95% of the words in the modified text were from the first and second 1000-word level in the British National Corpus, based on the Lextutor online program (Cobb, n. d.). According to teachers’ experience, these words were known to their students. Knowledge of 95% of these words was considered a reasonable percentage for adequate comprehension of the texts and allowed learners to concentrate on extracting the meaning of some new words while listening (van Zeeland & Schmitt, 2013).

In this research, the node word in a collocation was a verb with a noun as the collocate. The study focused on verb–noun collocations because they are challenging for L2 learners (Chan & Liou, 2005). Node words were low-frequency words. Sixteen collocations (exaggerate facts; accompany friends; abolish rules; decorate walls; purify water; rectify mistakes; scrub dirt; maintain business; uphold honesty; generate power; repel enemies; quench fires; modify rules; initiate activities; and execute plans) were consistent in each group. The collocations were pre-tested with 40 learners possessing similar language proficiency (i.e., achieving a score of 80–85 in the internal English proficiency test) and educational background but who were not participants in the main study. The pilot students were required to circle the correct collocate from five options based on the given node word. The five choices included a correct collocate, three distracters, and an “I don’t know” option. The students were told not to guess but to choose the “I don’t know” option if they were unsure or did not know the answer. Results showed that the target collocations were not known by the 40 learners. This is the most reliable assumption, as revealing target items to the participants would confound the findings. To make the test easier for the learners, 30 high-frequency collocations were mixed with the target collocations. Results showed that the pilot students did not know the meaning of the target collocations, which was the most reasonable assumption without asking participants in the main study to read the video texts and underline each unknown collocation; doing so would have contaminated the study findings. Collocations were inserted at appropriate places in the four episodes with the intention of maintaining a reasonable distribution.

An English native teacher read the story aloud while watching the original video. The spoken speed was in line with the caption speed, 90 words per minute, a safe rate for caption display in young learners’ television programs (Tyler et al., 2009). Wondershare Filmora, an online software, was used to edit a new audio file created by the native speaker to replace the original audio track. Captions presented a synchronized verbatim text of the soundtrack. Full or keyword captions for the video were added via MAGpie, a free online program. The keyword captions represented a single word (e.g., chatterbox) or a maximum of three consecutive words (e.g., exaggerate facts). Keywords were words essential for learners to identify the key meanings in sentences. The keywords represented approximately 19.6% of all words (i.e., 812 out of 4022 words). Target collocations were included in the keywords to make conditions comparable to the full-captioning group. Similar to full captions, keywords were centered on the captioning line. The presentation time for keywords was about 1 s, but the exact time was adjusted accordingly, as it was dependent on whether the keyword was a single-word item or multi-word unit. A joint discussion session was held between the six teachers. During the session, six teachers watched and discussed the videos. The teachers stated that the difficulty level, topic, and image functionality of videos were appropriate for the target students. Although the video images supported the content, the teachers stated that the images may not be highly explicit to determine meaning. Figs. 1 and 2 present screenshots of full captions and keyword captions.

3.3.2. An interactive advance-organizer strategy

The advance-organizer strategy in the present study refers to a teaching strategy that helps L2 learners integrate prior knowledge and organize their thoughts and ideas for improved input comprehension by exercising given ‘known’ information in videos. The advance-organizer strategy was assumed to potentially strengthen learners’ associations between pictorial and linguistic information, which could activate their verbal and pictorial cognitive systems to ensure meaningful learning. Information provided in the advance organizer was a brief synopsis about a video episode that learners would view. According to Herron (1994), the information synopsis served as a reading framework, allowing learners to link prior knowledge (content knowledge and/or linguistic knowledge) with upcoming events in the video episode for better comprehension. Consistent with previous studies (e.g., Li, 2014), authentic pictures taken directly from videos can provide accurate clues for learners to integrate video content with previously learned knowledge and facilitate learning progress.

The advance-organizer strategy used in this study contained 10 episodic pictures from each video. The strategy was designed to resemble a test in which EFL students’ attention would be focused on completing the task. Information presented using this strategy could also function as a ‘TV Guide’, and the students could understand upcoming information contained in the videos. Thus, this synopsis-viewing orientation may motivate learners to watch the videos and provide them prior knowledge (e.g., content knowledge and linguistic knowledge) to integrate into upcoming events in the videos for better comprehension. In this study, each picture was accompanied by a two-to-three-line English caption. All episodic pictures from four videos contained important information from the episodes, and some pictures included the target collocations. The six English teachers who assisted with this study chose the 10 pictures. These images were expected to help participants obtain the primary information. As indicated in Fig. 3, captions were displayed for learners on the top half of the screen, and pictures were displayed on the bottom half. Each page included four episodic pictures and four corresponding captions. Participants were instructed to choose one option (A, B, C, or D) from the right side of each picture. The selected option would appear in the box. If the selected option did not match the picture, an icon showing a sad face
popped up on the computer screen, and learners were prompted to try again. Learners sometimes took several attempts before completing a page successfully. An icon showing a happy face would pop up on the screen when a correct answer was selected. The learners then clicked on the message “GO ON TO THE NEXT PAGE” to continue to the next page. Learners could also listen to an audio recording of each line through their headsets while viewing the corresponding picture. This interactive advance-organizer activity ended when learners clicked on the ‘Submit’ icon after completing the last picture.

3.4. Measures

The post-test administered after the treatment conditions included four parts. The sequence of the four test parts proceeded as productive knowledge of form, receptive knowledge of form, productive knowledge of meaning, and receptive knowledge of
meaning. The test measured multiple aspects of collocational knowledge, which is more effective for accurately assessing learning outcomes compared to measuring a single aspect. The test was adapted from material by Webb et al. (2013). Different from their research that used a pencil-and-paper format, a computerized test was administered in this study.

In the first section, which measured productive knowledge of form, learners were required to write collocates based on a given word. Learners were also instructed to write the correct verb that had been encountered in the stories for the given noun. In the following example, participants were asked to supply the target collocate *exaggerate* beside the word *facts*:

___ facts

The second part was a multiple-choice test to measure receptive knowledge of form. Each item included five options: the correct collocate, three plausible distracters, and one “I don't know” option to reduce wild guessing. For example,

___ facts

A. determine B. exaggerate C. imagine D. decorate E. I don't know

Fig. 3. A screenshot of the interactive advance-organizer strategy.
The third part measured productive knowledge of meaning and was presented in a translation format; learners used L1 meanings to write L2 collocations. This test was intended to identify whether learners had incidentally learned the form and meaning of the collocations. For instance, learners were asked to write the collocation exaggerate facts beside its L1 translation:

Exaggerate the facts ______

The fourth part measured receptive knowledge of meaning in a receptive translation format. This test included L2 collocations as cues, and learners were instructed to write down the L1 meanings. In the following example, learners were asked to write the L1 meaning for the collocation exaggerate facts:

exaggerate facts ______

All test items were assigned one point for a correct answer and zero points for an incorrect answer. The learners had to spell the target collocation correctly to receive one point. Two teachers, who were not teaching any of the six groups, rated the tests independently. Any answers on which the raters could not reach a consensus required the involvement of a third rater, in which case the final decision was based on majority opinion. The two raters reached full consensus when marking the four test parts.

Several aspects of this test warrant consideration. First, the sequence of the four test parts was intended to decrease hints the learners may have received from preceding sections. Second, the order in which collocations appeared varied across the four test parts to reduce the possibility of a learning effect due to exposure to test information. Third, another set of 16 collocations selected randomly from the first 1000-word level was added randomly and mixed with test words to prevent repeated occurrence of target words from leading the participants to pay particular attention to the target collocations. Finally, learners could not return to test items after clicking “next item” in the computerized test. The time for the test, as suggested by the teachers, was 60 min. Cronbach's alpha for the four test parts ranged from 0.78 to 0.83, indicating a sound reliability level.

3.5. Procedures

Each school approved this study. Participants and their parents signed consent forms and were informed that learners would be required to watch English video clips and complete several exercises, although their performance would not affect their normal studies. However, they were not informed of the nature of the study (i.e., testing learning outcomes of collocations using videos and an advance-organizer technique), which allowed the research to be focused on incidental learning (Hulstijn, 2013). As an incentive, students received a meal coupon for participating in the study.

This study was conducted during special class periods on a Saturday in a computer laboratory. All students were provided a brief introduction to the videos and made aware of the need to focus on the content. Each student was assigned to a computer to work individually. Learners who received the advance-organizer strategy proceeded through an additional tutorial about how to work with the advance-organizer activity. The learners first completed the online advance-organizer activity and then viewed videos. Each videoclip was accessible via a hyperlink, which students were allowed to watch only once. Each teacher, who was familiarized with the research design through a joint session, served as a monitor for each group and ensured the learners did not have trouble watching videos and that the procedure was followed appropriately. The learners immediately proceeded to the test once they finished watching the videos. The experiment took approximately 2 h to complete.

3.6. Data analysis

The effects of captioned videos and advance organizers on incidental learning of collocations were measured using a two-way multivariate ANOVA (MANOVA) as data were normally distributed (the p value was larger than 0.05 for the Shapiro–Wilk test). Dependent variables were the four dimensions of the vocabulary test, which were significantly correlated with values ranging from 0.37 to 0.39, suggesting the assumption of running MANOVA was met. The two independent variables were “advance organizers”, which included two categories (i.e., “watching videos with advance organizers” and “watching videos without advance organizers”) and “captioned videos”, with three categories (i.e., “full captioning”, “keyword captioning”, and “no captions”). The significance level was set at 0.05, with a Bonferroni correction to reduce the Type I error rate.

Table 1
Descriptive statistics of four vocabulary test parts.

<table>
<thead>
<tr>
<th>Advance-organizer (AO) technique</th>
<th>Type of captioning</th>
<th>n</th>
<th>M</th>
<th>SD</th>
<th>M</th>
<th>SD</th>
<th>M</th>
<th>SD</th>
<th>M</th>
<th>SD</th>
</tr>
</thead>
<tbody>
<tr>
<td>Presence of AO</td>
<td>FC</td>
<td>59</td>
<td>5.36</td>
<td>1.27</td>
<td>9.76</td>
<td>1.04</td>
<td>7.12</td>
<td>1.01</td>
<td>12.37</td>
<td>1.16</td>
</tr>
<tr>
<td>Presence of AO</td>
<td>KC</td>
<td>54</td>
<td>3.46</td>
<td>1.13</td>
<td>7.01</td>
<td>1.16</td>
<td>5.01</td>
<td>1.24</td>
<td>9.89</td>
<td>1.05</td>
</tr>
<tr>
<td>Presence of AO</td>
<td>NC</td>
<td>58</td>
<td>1.28</td>
<td>.97</td>
<td>5.03</td>
<td>1.17</td>
<td>3.14</td>
<td>1.18</td>
<td>7.02</td>
<td>1.06</td>
</tr>
<tr>
<td>Absence of AO</td>
<td>FC</td>
<td>67</td>
<td>3.26</td>
<td>1.07</td>
<td>6.88</td>
<td>1.26</td>
<td>4.99</td>
<td>1.01</td>
<td>9.87</td>
<td>1.17</td>
</tr>
<tr>
<td>Absence of AO</td>
<td>KC</td>
<td>61</td>
<td>1.38</td>
<td>.98</td>
<td>4.89</td>
<td>1.11</td>
<td>3.12</td>
<td>.91</td>
<td>7.03</td>
<td>1.05</td>
</tr>
<tr>
<td>Absence of AO</td>
<td>NC</td>
<td>62</td>
<td>.61</td>
<td>.74</td>
<td>2.03</td>
<td>.91</td>
<td>1.15</td>
<td>.86</td>
<td>3.88</td>
<td>1.09</td>
</tr>
</tbody>
</table>

Note. The maximum score for each test part is 16 points. FC = Full captioning, KC = keyword captioning, NC = No captioning.
indicating that use of an advance-organizer strategy yielded significant improvements in learners’ receptive meaning compared to the absence of this technique. Similarly, as Table 5 shows, the main effect of captions on incidental learning of collocations, post-hoc pairwise comparisons between using an advance organizer or not in each caption condition. Results are presented in Table 6, indicating that use of an advance-organizer strategy yielded significantly better results than the absence of an advance-organizer technique. Learners thus benefited from the advance-organizer technique irrespective of whether they encountered collocations in videos with full captions, keyword captions, or no captions. One exception was that using the advance-organizer strategy did not lead to significantly improved performance in learning productive form under the no-caption condition (p > .05), implying that when students encountered collocations in videos without captions, advance-organizer techniques did not engender better results in recalling productive form than in the absence of this technique.

Similarly, as Table 5 shows the main effects of captioned videos on incidental learning of collocations, post-hoc pairwise analysis

### 4. Results

Table 1 presents descriptive statistics for each test section. Under each captioning condition, the use of the advance-organizer strategy yielded higher scores than the absence of the advance-organizer technique. In addition, under each advance-organizer condition, full captioning led to higher scores, followed by keyword captioning and no captions. Among the six groups, the FC + AO group achieved the best results on the four test parts (productive form: 5.36; receptive form: 9.76; productive meaning: 7.12; receptive meaning: 12.37).

Table 2 presents the results of Box's test of equality of covariance matrices, which measured the null hypothesis that the observed covariance matrices of the dependent variables were equal across groups. The p value was larger than the alpha level (0.05), indicating that the matrices were equal; hence, the assumption of homogeneity was met, suggesting the appropriateness of reporting the Wilks’ lambda results.

Table 3 presents the multivariate test results. The Wilks’ lambda results revealed differences between the group means on dependent variables. Based on significance values of the F-ratios, findings indicated a statistically significant effect of the advance-organizer intervention on learners’ incidental learning of collocations [F (4, 231) = 165.814, p < .001; Wilks’ Λ = 0.258, partial η² = 0.468]. A statistically significant effect also appeared for captioned videos on incidental learning of collocations [F (8, 462) = 111.905, p < .001; Wilks’ Λ = 0.116, partial η² = 0.660]. A statistically significant interaction effect was found between advance organizers and types of captioned videos on the combined dependent variables [F (8, 462) = 82.026, p < .05; Wilks’ Λ = 0.053, partial η² = 0.062].

Table 4 presents the results of Levene’s test of equality of variances for each dependent variable, which evaluated the null hypothesis that the error variance of the dependent variable was equal across groups. Findings showed that all dependent variables were non-significant (p > .05). These results support the reliability of univariate tests (ANOVA) conducted following MANOVA.

Table 5 presents results of univariate tests (ANOVA summary) conducted after MANOVA for the dependent variables. Findings demonstrate the main effects of advance-organizer conditions on productive form (p < .001, F = 29.128, partial η² = 0.418), receptive form (p < .001, F = 46.175, partial η² = 0.468), productive meaning (p < .001, F = 55.188, partial η² = 0.563), and receptive meaning (p < .001, F = 36.167, partial η² = 0.475). Results also suggest main effects of captioned video conditions on productive form (p < .001, F = 45.465, partial η² = 0.545), receptive form (p < .001, F = 36.576, partial η² = 0.467), productive meaning (p < .001, F = 36.587, partial η² = 0.485), and receptive meaning (p < .001, F = 47.686, partial η² = 0.486).

Table 6 lists the main effects of advance-organizer conditions on the four dependent variables; the next step was to determine post-hoc pairwise comparisons between using an advance organizer or not in each caption condition. Results are presented in Table 6, indicating that use of an advance-organizer strategy yielded significantly better results than the absence of an advance-organizer technique. Learners thus benefited from the advance-organizer technique irrespective of whether they encountered collocations in videos with full captions, keyword captions, or no captions. One exception was that using the advance-organizer strategy did not lead to significantly improved performance in learning productive form under the no-caption condition (p > .05), implying that when students encountered collocations in videos without captions, advance-organizer techniques did not engender better results in recalling productive form than in the absence of this technique.

Similarly, as Table 5 shows the main effects of captioned videos on incidental learning of collocations, post-hoc pairwise analysis

<table>
<thead>
<tr>
<th>Table 2</th>
<th>Results of Box’s M tests.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Box’s Test of Equality of Covariance Matricesa</td>
<td></td>
</tr>
<tr>
<td>Box’s M</td>
<td>344.055</td>
</tr>
<tr>
<td>F</td>
<td>5.583</td>
</tr>
<tr>
<td>df1</td>
<td>50</td>
</tr>
<tr>
<td>df2</td>
<td>100497.306</td>
</tr>
<tr>
<td>p</td>
<td>0.651</td>
</tr>
</tbody>
</table>

  a Design: Captioned videos + advance organizers + advance * organizers.

<table>
<thead>
<tr>
<th>Table 3</th>
<th>Multivariate test results.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Effect</td>
<td>Value</td>
</tr>
<tr>
<td>Intercept</td>
<td>Wilks’ Lambda</td>
</tr>
<tr>
<td>Advance organizers</td>
<td>Wilks’ Lambda</td>
</tr>
<tr>
<td>Captioned videos</td>
<td>Wilks’ Lambda</td>
</tr>
<tr>
<td>Advance organizers * Captioned videos</td>
<td>Wilks’ Lambda</td>
</tr>
</tbody>
</table>

  a Design: Intercept + advance organizers + captioned videos + advance organizers * captioned videos.
  b Exact statistic.
was carried out to assess the effects of different captions in each condition with or without the advance-organizer strategy. Results of pairwise comparisons are presented in Table 7. Overall, the $F$ and $p$ values reveal significant differences between the three captioning conditions under each advance-organizer condition; full captioning contributed to better incidental learning of collocations than keyword captioning ($p < .001$), and keyword captioning contributed to better incidental learning of collocations than no captions ($p < .001$). The only exception involved a non-significant difference in learning productive form when comparing the keyword-captioning and no-caption conditions without an advance-organizer strategy ($p > .05$).

The next step was to explore how six conditions—i.e., the different combinations of advance-organizer technique and captioning type—would compare to each other. Table 8 shows the comparisons between all possible pairs of conditions. In terms of productive form, results showed that encountering collocations in FC + AO was more effective (significantly different) than encountering them

### Table 4
Results on Levene's test of equality of error variances.

<table>
<thead>
<tr>
<th>Source</th>
<th>Type III Sum of Squares</th>
<th>df1</th>
<th>df2</th>
<th>$F$</th>
<th>$p$</th>
<th>Partial Eta Squared</th>
</tr>
</thead>
<tbody>
<tr>
<td>Corrected Model</td>
<td>Productive form</td>
<td>812.671</td>
<td>5</td>
<td>162.534</td>
<td>18.468</td>
<td>0.000</td>
</tr>
<tr>
<td></td>
<td>Receptive form</td>
<td>537.544</td>
<td>5</td>
<td>107.502</td>
<td>32.857</td>
<td>0.000</td>
</tr>
<tr>
<td></td>
<td>Productive meaning</td>
<td>912.737</td>
<td>5</td>
<td>182.547</td>
<td>26.156</td>
<td>0.000</td>
</tr>
<tr>
<td>Advance organizers</td>
<td>Productive form</td>
<td>256.557</td>
<td>1</td>
<td>256.557</td>
<td>29.128</td>
<td>0.000</td>
</tr>
<tr>
<td></td>
<td>Receptive form</td>
<td>234.565</td>
<td>1</td>
<td>234.565</td>
<td>46.175</td>
<td>0.000</td>
</tr>
<tr>
<td></td>
<td>Productive meaning</td>
<td>367.856</td>
<td>1</td>
<td>367.856</td>
<td>55.188</td>
<td>0.000</td>
</tr>
<tr>
<td></td>
<td>Receptive meaning</td>
<td>334.018</td>
<td>1</td>
<td>334.018</td>
<td>36.167</td>
<td>0.000</td>
</tr>
<tr>
<td>Captioned videos</td>
<td>Productive form</td>
<td>535.557</td>
<td>2</td>
<td>267.778</td>
<td>45.465</td>
<td>0.000</td>
</tr>
<tr>
<td></td>
<td>Receptive form</td>
<td>554.635</td>
<td>2</td>
<td>277.317</td>
<td>36.576</td>
<td>0.000</td>
</tr>
<tr>
<td></td>
<td>Productive meaning</td>
<td>556.554</td>
<td>2</td>
<td>278.277</td>
<td>36.587</td>
<td>0.000</td>
</tr>
<tr>
<td></td>
<td>Receptive meaning</td>
<td>767.676</td>
<td>2</td>
<td>383.838</td>
<td>47.868</td>
<td>0.000</td>
</tr>
</tbody>
</table>

$a$ R Squared = .243 (Adjusted R Squared = .211).


$c$ R Squared = .268 (Adjusted R Squared = .212).

$d$ R Squared = .269 (Adjusted R Squared = .224).

### Table 5
Results on univariate tests that follow MANOVA.

#### Tests of Between-Subjects Effects

<table>
<thead>
<tr>
<th>Source</th>
<th>Type III Sum of Squares</th>
<th>df</th>
<th>Mean Square</th>
<th>$F$</th>
<th>$p$</th>
<th>Partial Eta Squared</th>
</tr>
</thead>
<tbody>
<tr>
<td>Corrected Model</td>
<td>Productive form</td>
<td>812.671</td>
<td>5</td>
<td>162.534</td>
<td>18.468</td>
<td>0.000</td>
</tr>
<tr>
<td></td>
<td>Receptive form</td>
<td>537.544</td>
<td>5</td>
<td>107.502</td>
<td>32.857</td>
<td>0.000</td>
</tr>
<tr>
<td></td>
<td>Productive meaning</td>
<td>912.737</td>
<td>5</td>
<td>182.547</td>
<td>26.156</td>
<td>0.000</td>
</tr>
<tr>
<td>Advance organizers</td>
<td>Productive form</td>
<td>256.557</td>
<td>1</td>
<td>256.557</td>
<td>29.128</td>
<td>0.000</td>
</tr>
<tr>
<td></td>
<td>Receptive form</td>
<td>234.565</td>
<td>1</td>
<td>234.565</td>
<td>46.175</td>
<td>0.000</td>
</tr>
<tr>
<td></td>
<td>Productive meaning</td>
<td>367.856</td>
<td>1</td>
<td>367.856</td>
<td>55.188</td>
<td>0.000</td>
</tr>
<tr>
<td></td>
<td>Receptive meaning</td>
<td>334.018</td>
<td>1</td>
<td>334.018</td>
<td>36.167</td>
<td>0.000</td>
</tr>
<tr>
<td>Captioned videos</td>
<td>Productive form</td>
<td>535.557</td>
<td>2</td>
<td>267.778</td>
<td>45.465</td>
<td>0.000</td>
</tr>
<tr>
<td></td>
<td>Receptive form</td>
<td>554.635</td>
<td>2</td>
<td>277.317</td>
<td>36.576</td>
<td>0.000</td>
</tr>
<tr>
<td></td>
<td>Productive meaning</td>
<td>556.554</td>
<td>2</td>
<td>278.277</td>
<td>36.587</td>
<td>0.000</td>
</tr>
<tr>
<td></td>
<td>Receptive meaning</td>
<td>767.676</td>
<td>2</td>
<td>383.838</td>
<td>47.868</td>
<td>0.000</td>
</tr>
</tbody>
</table>

$a$ R Squared = .243 (Adjusted R Squared = .211).


$c$ R Squared = .268 (Adjusted R Squared = .212).

$d$ R Squared = .269 (Adjusted R Squared = .224).

### Table 6
Results on comparing the advance organizer conditions in each caption condition.

#### Pairwise Comparisons

<table>
<thead>
<tr>
<th>Dependent Variables</th>
<th>Mean Difference (I-J)</th>
<th>$F$</th>
<th>$p$</th>
<th>Partial Eta Squared</th>
</tr>
</thead>
<tbody>
<tr>
<td>Productive form</td>
<td>FC With AO Without AO</td>
<td>2.10**</td>
<td>10.189</td>
<td>.000</td>
</tr>
<tr>
<td></td>
<td>KC With AO Without AO</td>
<td>2.08**</td>
<td>12.149</td>
<td>.000</td>
</tr>
<tr>
<td></td>
<td>NC With AO Without AO</td>
<td>0.67</td>
<td>91</td>
<td>.664</td>
</tr>
<tr>
<td>Receptive form</td>
<td>FC With AO Without AO</td>
<td>2.88**</td>
<td>16.781</td>
<td>.000</td>
</tr>
<tr>
<td></td>
<td>KC With AO Without AO</td>
<td>2.12**</td>
<td>16.212</td>
<td>.000</td>
</tr>
<tr>
<td></td>
<td>NC With AO Without AO</td>
<td>2.00**</td>
<td>18.171</td>
<td>.000</td>
</tr>
<tr>
<td>Passive meaning</td>
<td>FC With AO Without AO</td>
<td>2.13**</td>
<td>18.164</td>
<td>.000</td>
</tr>
<tr>
<td></td>
<td>KC With AO Without AO</td>
<td>1.89**</td>
<td>10.126</td>
<td>.002</td>
</tr>
<tr>
<td></td>
<td>NC With AO Without AO</td>
<td>1.99*</td>
<td>10.194</td>
<td>.002</td>
</tr>
<tr>
<td>Receptive meaning</td>
<td>FC With AO Without AO</td>
<td>2.50**</td>
<td>21.218</td>
<td>.000</td>
</tr>
<tr>
<td></td>
<td>KC With AO Without AO</td>
<td>2.86**</td>
<td>22.258</td>
<td>.000</td>
</tr>
<tr>
<td></td>
<td>NC With AO Without AO</td>
<td>3.14**</td>
<td>25.264</td>
<td>.000</td>
</tr>
</tbody>
</table>

$a$ The mean difference is significant at the 0.05 level.

$**$ The mean difference is significant at the 0.001 level.

Note. FC = Full captions, KC = keyword captions, NC = no captions, AO = Advance organizers.

$^b$ Adjustment for multiple comparisons: Bonferroni.
in full captioning, keyword captioning, and no words captioning without receiving advance-organize technique. Encountering words in the KC + AO was not found to be more effective than FC, but was found to be more effective than encountering the words in the KC and NC conditions. Likewise, there was a significant difference between FC and NC + AO. However, NC + AO was not observed to be more effective than KC and NC conditions. The results were slightly different for receptive form. Combinations that were not significantly different from each other included: KC + AO vs. FC; NC + AO vs. FC; and NC + AO vs. KC. In terms of productive meaning, combinations (KC + AO vs. FC; and NC + AO vs. KC) were not significantly different. Slight differences were found in the learning of receptive meaning for which combinations (KC + AO vs. FC; NC + AO vs. KC) were not significantly different. Overall, these patterns suggest that both captioning type and advance-organizer strategy had a pronounced effect on the incidental learning of collocations. In particular, the combination of using advance-organizer strategy in the full captioning condition yielded significantly better results than any of the other five combinations. This was consistent for learning the four dimensions of collocation knowledge.

5. Discussion

Findings highlight the effect of an advance-organizer strategy on incidental learning of collocations for L2 learners. These results correspond with those of research on the facilitative role of an advance organizer in enhancing L2 learning (Chung & Huang, 1998;
Herron, 1994; Li, 2014), reinforcing the assertion that an advance-organizer strategy is effective in enhancing incidental learning of collocations. The assistance provided by the advance-organizer strategy may help L2 learners activate prior knowledge and synthesize important information, which may in turn reduce learners' cognitive processing load for upcoming video information to promote meaningful learning and increased performance when learning different dimensions of collocations. As argued by Mayer (2001), the advance-organizer strategy facilitated L2 learners to outline, organize, and sequence main ideas from watching videos. The advance-organizer strategy in the current study likely functioned as a cognitive tool and presented balanced organization of information to afford learners opportunities to discern information in the videos. Tang (1992) argued that organizing new information enhances information processing in working memory and facilitates comprehension necessary for constructing a form–meaning link. Hence, L2 learners might take advantage of organized information and thus enhance their ability to notice verb + noun collocations from language input.

The significant difference in immediate post-test scores across the three captioning groups (i.e., full captions, keyword captions, and no captions) also demonstrated the effects of captioned videos on incidental learning of collocations. Fully captioned videos led primary school L2 students to exhibit improved incidental learning of verb–noun lexical collocation knowledge to the greatest extent, followed by videos with keyword captioning and no captioning. Vanderplank (2016) argued that captioning combines different visuals (pictures and words) with auditory stimuli, which is effective for L2 learning. Findings on positive learning of collocations has also demonstrated the benefits of captions for L2 learning (e.g., Montero Perez et al., 2018, 2013; Peters et al., 2016). The outcome of this study is in line with research revealing that full captioning produced greater gains than keyword captioning and no captioning (Montero Perez et al., 2013; Teng, 2019a). However, Montero Perez et al. (2018) found that keyword captioning was more facilitative for incidental vocabulary learning. Peters et al. (2016) asserted that learners' vocabulary size may be a substantial factor affecting L2 learners' performance in incidental vocabulary learning from watching captioned videos. In the present study, L2 learners' English proficiency level, including reading, writing, listening, and speaking, was homogeneous. In addition, participants achieved scores of 80–85 out of 100 points on an English language assessment and may have exhibited better language proficiency levels than other L2 learners from the same background. As argued by Vanderplank (2010) and Teng (2019b), captions seemed to be better suited for learners with higher language proficiency. Although the aim of the present study was not to explore the complex relationship between English proficiency level and captioning types, these findings highlight the need for more studies to clarify the role of English proficiency and other variables (e.g., types of captioning) in learning words incidentally through audiovisual input. When learners' English proficiency was quite low, learning occurred rarely or not at all as learners likely found their cognitive processing limited when handling full and complete captions at a fast pace.

Another explanation for this result could be that full captioning, compared to keyword captioning, might have provided more information in this study to enable learners to process linguistic and content messages. Primary school L2 learners' enhanced learning of collocations through fully captioned videos seemed plausible, as these students were mostly acquiring L2 vocabulary in a classroom setting and depended on printed materials, compared to native English students for whom learning collocations might be easier due to having more authentic language exposure outside the classroom (Wible, 2008). Consistent with Teng (2019a) study, full captions possibly helped L2 learners construct links between spoken words and printed words and draw upon background vocabulary knowledge and comprehension strategies to discern potential connections between word form and meaning. However, a comparison of learning gains through research studies should be interpreted cautiously due to the diverse quantity of target items, differences between learning single words or collocations, the variety of information embedded in captions, and the lengths of the interventions.

Results showed a significant interaction effect of full captions and advance organizers on incidental learning of collocations. As indicated in Table 8, learners who used an advance-organizer strategy in the full-caption condition significantly outperformed learners in other groups. In the post-test (Table 1), out of 16 points, these learners earned an average score of 12.37 and 9.76 for meanings and forms of target collocations and an average score of 7.12 and 5.36 for meanings and forms of collocations. This finding suggests that the L2 young learners obtained substantial collocation knowledge, especially in recognizing form–meaning links. In a recent study (Montero Perez et al., 2018), learners' performance in recognizing word form and meaning after watching full captioned videos was 9 and 10.29 out of 18 points. The least-beneficial learning effect was on word-meaning recall, on which students scored only 2.88 out of 18 points. Information presented using the advance-organizer strategy may help learners activate their background knowledge and develop a holistic picture of target video materials by discerning upcoming details of the video plot (Li, 2014). In addition, the advance organizer condition may have resulted in more exposures to the collocations. Corroborated previous studies (Teng, 2019c; Teng, 2019d; Webb et al., 2013), the repeated exposure may lead to more robust incidental vocabulary learning outcomes. Learners may be engaged with unfamiliar collocations rather than listening to and reading the text to identify the main idea. Learners' encouraging performance in this study revealed the following: a) it is feasible to focus on incidental learning of collocations, a more challenging component than single words; and b) young learners, who may have limited cognitive processing compared to L2 adult learners, can be facilitated to learn collocations through captions and advance-organizers.

Results suggest that receptive knowledge of collocations was enhanced more easily than recall knowledge, consistent with studies on collocations when reading printed texts (e.g., Pellicer-Sánchez, 2017), online reading (Teng, 2018b) or reading and listening to a modified graded reader (Webb et al., 2013). Knowledge acquisition of collocations appeared to occur at a rate similar to the incremental process of learning single lexical items (e.g., Webb & Kagimoto, 2011). For example, the nature of learning the productive form of a collocation may be more difficult than learning receptive knowledge of lexical mastery. Encountering new words using fully captioned videos and the advance-organizer strategy may boost learners' capacity to consolidate collocational knowledge and help them select information about form and meaning for new collocates to retain knowledge about which words co-occur frequently. Partial gains in learning collocations may reflect inherent difficulties in learning collocations, insufficient exposure to new collocations, and the short period of time for which learners were presented with audiovisual input. As Barcroft (2002) asserted, L2
learners' limited cognitive resources may lead them to overemphasize one aspect (e.g., form) while neglecting information about other aspects (e.g., meaning). Hence, similar to the learning of single words, L2 learners may be able to learn more automatic collocation knowledge aspects (e.g., form recognition) but may not process information for less automatic aspects (e.g., meaning recall). Recalling meaning for collocates of unknown words was found to be a nonlinear and incremental process through which collocation-focused instructional procedures (e.g., via advance-organizer strategies prior to viewing fully captioned videos) could enrich learners’ mental lexicon.

6. Concluding remarks

Overall, results from this empirical study support the use of advance organizers and watching captioned videos to facilitate primary school L2 learners' collocation knowledge and can be considered an alternative approach to teaching collocations in a context where L2 learners typically receive input through printed texts. This study is innovative given the increasing importance of new multimedia technologies in L2 classroom settings and the growing value in applying technology in vocabulary learning. Although the number of collocations learned was partial and restricted, this study provides insight into primary school L2 learners' progression from no knowledge of an item to form recognition, meaning recognition, and finally to meaning recall by watching captioned videos and employing an advance-organizer strategy.

Results should be interpreted with caution due to certain limitations. First, participants in this study were advanced primary school Chinese students. The learners may focus on the collocations, and thus gain more scores in four vocabulary tests. Thus the findings may not be generalizable to learners with low English proficiency in a reading-texts-dependent English learning setting or in other contexts with learners whose first language is not Chinese. Second, the types of advance organizers used may temper the generalizability of these findings. Third, prior research included different types of videos for L2 learning, such as television reports (Montero Perez et al., 2013), television programs (Rodgers & Webb, 2017), documentary videos (Markham, Peter, & McCarthy, 2001), and storytelling videos (the present study). Subsequent studies should assess whether different types of audiovisual input generate different effects. Fourth, this study did not examine performance differences between genders or learners with varying levels of language proficiency, presenting opportunities for scholars to replicate this study with learners from other contexts. Finally, the complexity of collocations (i.e., varied relationships between nodes and collocates, such as adjective–noun, and different levels of transparency) implies that the results of a study using only verb–object collocations may not be generalizable to the learning and teaching of other types of collocations; hence, additional research is needed in this vein. In addition, this study used the same video images as advance-organizers. Learners may have generalized the meaning of the video images from the full captions and the practice effect may occur.

Despite these limitations, this study is innovative in investigating different combinations of captioning types and advance organizers. Future studies should address whether interactive advance-organizer strategies help L2 learners when interventions are administered during or after watching videos and whether differential effects emerge when learning collocations based on different organizer-aided interventions. Given the limited time for this experiment, future studies are also needed for understanding the long-term learning effect with the same method on incidental L2 collocation learning. Future studies should also evaluate the correlations based on connecting learning style (or preferred multimedia type) to the captioned videos and organizer strategies. The teaching effectiveness of collocations could be expanded by viewing captioned videos and completing an advance-organizer strategy outside or within the classroom, providing important pedagogical implications. First, teaching of collocations tends to be neglected in L2 vocabulary lessons, and L2 learners may be weak in collocation use. This study's findings support claims of using captioned videos in L2 learning and the belief that teachers should introduce captioned videos as a supplement to teaching lexical chunks explicitly. With this approach, L2 learners may attend to multiword units outside the classroom and learn more collocations incidentally from multimedia materials. Second, an organizer-aided activity can be introduced before viewing captioned videos to help L2 learners activate and process prior and upcoming mental knowledge in videos. An organizer-aided activity may encourage L2 learners to employ strategies to discern and strengthen what has or has not been learned, thus enabling learners to simultaneously process visual images and spoken text without sensory overload. Although captioned videos may highlight pictures and words, an interactive organizer-aided activity may assist L2 learners in applying strategies to identify useful information stored in working memory for better comprehension of information and more effective incidental learning of collocations. Finally, contributing to previous studies in delineating the partial gains in learning phrasal verbs (e.g., Teng, 2017), the partial gains of learning collocations, particularly limited gains in recalling collocation meaning, suggest that learning collocations remains challenging in L2 vocabulary acquisition. Future studies can extend the present work to develop knowledge related to anticipated trajectories when learning specific components (i.e., form, meaning, and use) of collocations.

Conflicts of interest

The author state that there is no conflict of interest for this article.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.compedu.2019.103655.
References


(Mark) Feng Teng is a language teacher educator with extensive teaching experience in China. His professional interests include metacognition and writing, L2 vocabulary development, and identity research. He has published widely in international flagship journals, including Language Teaching Research, Applied Linguistics, TESOL Quarterly, Applied Linguistics Review, Computer Assisted Language Learning, and others. He is editing several special issues for international journals.