the jaltcalljournal

ISSN 1832-4215
Vol. 14, No.2 Pages 119–142

Incidental professional vocabulary acquisition of EFL business learners: Effect of captioned video with glosses

Hui-Tzu Hsu

National Taiwan Normal University, Taiwan
lindahsu85@gmail.com

Use of captioned video in classrooms has gained considerable attention in the second and foreign language learning. However, the effect of application of captioned video embedded with glosses on incidental vocabulary enhancement has not been explored. This study aims to examine the effect of video captions with glosses on EFL students' incidental business vocabulary acquisition; 50 students from a college of management served as participants. A pretest was adopted to ensure participants lacked familiarity with the target vocabulary. All participants watched three video clips presented in three modes (noncaptioned, captioned, and caption-gloss modes). After each mode session, all participants took an immediate posttest and a 3-week-delayed posttest. Following the final posttest, the participants completed a questionnaire. The findings revealed that the caption-gloss mode significantly outperformed the other 2 modes in both the immediate and 3-week-delayed posttests. Retention of the target business words did not significantly decline at the 3-week-delayed posttest. Therefore, glosses in the captioned video improved the participants' short- and long-term incidental business vocabulary retention. The participants also provided positive feedback regarding the efficacy of the caption-gloss mode for incidental business word acquisition. Pedagogical implications regarding use of captioned video with glosses for incidental professional vocabulary acquisition are discussed.

Keywords: Captioned video; Glosses; Incidental vocabulary learning; Professional vocabulary

Regular Paper

Introduction

Increasing emphasis has been placed on the effectiveness of multimedia in language teaching and learning. Harji, Woods, and Alavi (2010) stated that multimedia technology such as TV, computers, networks, email, videos, compact disc read-only memories, and interactive multimedia may be adopted to integrate authentic target language materials into a language classroom environment. Multimedia can be used to provide rich and authentic inputs for comprehension, and use of these materials exhibits a positive effect on language education (Brett, 1995; Dwyer, 1978; Magasic, 2017; Vanderplank, 2010). Specifically, multimedia can be used for vocabulary learning. According to Burston (2015), multimedia messaging services, including those involving text, pictures, sound, and video, can be applied in alternative methods for vocabulary learning. Multimedia annotations have received attention in research on vocabulary learning, and textual definitions with pictures have outperformed other tested combinations (Chun & Pass, 1996; Plass, Chun, Mayer, & Leutner, 1998; Shahrokni, 2009; Yoshii, 2006). Findings from research on multimedia annotations are supported by Paivio's (1986, 1991, 2007) dual coding theory, which suggests that people process combinations of imagery and verbal information relatively efficiently. However, conflicting results on the effects of text with pictures versus text with dynamic images have drawn substantial attention (Sydorenko, 2010). Multimedia annotations that combine video and text have been reported as more effective for vocabulary learning than annotations comprising pictures and text (Al-Seghayer, 2001).

Video with on-screen text can be beneficial for vocabulary learning. In this context, on-screen text may be categorized as that for native-language (L1) or second language (L2) learning. Text types comprise L1 subtitles (L2 video, L1 on-screen text), reversed subtitles (L1 video, L2 text), and captioning (L2 video, L2 text). Numerous studies have demonstrated that captioned videos benefit learners' vocabulary development (Baltova, 1999; Borras & Lafayette, 1994; Garza, 1991; Guillory, 1998; Neuman & Koskinen, 1992; Sydorenko, 2010; Vanderplank, 2013). Accordingly, an increasing number of language teachers and researchers have considered investigating the simultaneous presentation of L2 text and video (i.e., bimodal input), and the results of studies regarding the effectiveness of captioning for vocabulary learning have been promising (Danan, 1992; Montero Perez, Noortgate, & Desmet, 2013; Neuman & Koskinen, 1992; Sydorenko, 2010; Winke, Gass, & Sydorenko, 2010).

Captioned video and glosses can be used to draw the attention of L2 learners to unfamiliar vocabulary, thereby enabling incidental vocabulary acquisition. Hulstijn (2001) noted that researchers have identified attention as a crucial for vocabulary learning. Allocation of sufficient attention to an unfamiliar item is a critical initial step in the vocabulary acquisition process. Empirical evidence indicates that fully captioned video content enhances vocabulary learning more than noncaptioned video content (Danan, 1992; Markham, 1999; Montero Perez, Peters, Clarebout, & Desmet, 2014; Sydorenko, 2010). Winke et al. (2010) noted that full captioning helps L2 learners to isolate unknown vocabulary items, because the captions draw the learners' attention to the items. The critical role of attention for language acquisition has been widely verified in studies on second language acquisition (SLA) (Gass, 1999; Skehan, 1998). Hulstijn (2001) emphasized that tasks that draw learners' attention to new vocabulary are critical to learning progress.

Numerous studies have compared the effectiveness of various caption formats (no captions, full captions, keyword captions, full captions with highlighted keywords) for

vocabulary acquisition (Laufer, 2010; Montero Perez et al., 2014, Montero Perez, Peters, & Desmet, 2015; Peters & Montero Perez, 2015; Sydorenko, 2010). In addition to these caption-related formats, glosses can enhance vocabulary learning (Ko, 2005; Sanders, 2002) and draw learners' attention to the form and meaning of annotated words (Marzban, 2011). Montero Perez, Peters, and Desmet (2018) reported that their study was the first to investigate the use and effects of glossed keyword captions. The keywords in their study were French target words. Each glossed keyword appeared in isolation and was linked to its corresponding L1 context-bound translation. By contrast, the present study focused on learners' incidental acquisition of professional English vocabulary. Thus, the video caption with glosses employed to stimulate learners' incidental acquisition of target words differed from that described in the study by Montero Perez et al. (2018). In the present study, target professional English words were embedded with glosses in captions. Specifically, each glossed target word was marked within a caption rather than appearing in isolation, and was accompanied by an L1 translation and an L2 definition. The present study investigated whether learners could incidentally acquire professional English words when they watched a captioned video with the words glossed according to the described model.

Research has indicated that captioned videos enhance vocabulary acquisition. However, the relevant studies have not resolved two pressing questions. First, the use of captions with glosses for vocabulary learning has been somewhat neglected. As described, attention is critical to incidental acquisition of vocabulary (Hulstijn, 2001). However, no study has examined the effect of caption glosses on L2 learners' attention to target vocabulary and incidental vocabulary learning. In other words, researchers have regarded vocabulary learning as a by-product of listening for meaning (Gass, 1999). Captions with glosses as an input have received little attention in the literature on multimedia with glosses for vocabulary learning. Second, most related studies have focused on the effectiveness of captions for enhancing knowledge of general English vocabulary, but few empirical studies have investigated the effects of captions with glosses on professional vocabulary acquisition within the domain of English for specific purposes (ESP). Professional vocabulary is defined as technical, specialized vocabulary (Harding, 2007). ESP courses mainly emphasize language use, language function, and terminology (AbdulGhani, 1993). ESP addresses the vast field of technical terminology (Chung & Nation, 2004; Fraser, 2006). The purpose of the present study was to address the gaps in the literature concerning multimedia glosses for incidental vocabulary learning. Specifically, the focus of this research was the effectiveness of captions with glosses in an ESP context. Three modes (noncaptioned, captioned, caption-gloss) were used for the investigation.

Literature review

Vocabulary learning with a multimedia glossary

Incidental vocabulary learning and the use of glosses have been the focus of numerous studies on enhancing L2 vocabulary knowledge and instruction. According to Hulstijn (1989) and Schmidt (1994), many words are "picked up" during listening and reading activities when the listener or reader's main goal is to comprehend the meaning of the language heard or read, rather than to learn new words. This "picking up" is regarded as incidental learning. Nation (1999) stressed the critical nature of incidental learning that occurs through message-focused activities. L2 language learners can improve their incidental

vocabulary acquisition by using glosses (Hulstijn, Hollander, & Greidanus, 1996). Among related methods, marginal vocabulary glosses may be used to enhance incidental vocabulary learning (Hulstijn, 1992; Jacobs, Dufon, & Fong, 1994; Watanabe, 1997). Numerous studies have demonstrated that glosses in printed materials and electronic glossaries can be used to stimulate incidental L2 vocabulary acquisition (Al-Seghayer, 2001; Groot, 2000; Hulstijn et al., 1996; Jacobs et al., 1994; Laufer & Hill, 2000). In one such study, researchers integrated multimedia annotations into digital materials to facilitate learners' vocabulary acquisition (Chun & Plass, 1996). Multimedia annotations can comprise text, pictures, videos, animations, and sounds.

In summary, relevant studies have reached two primary conclusions. First, learners who acquire vocabulary through multiple annotations outperform those who acquire vocabulary through materials with single annotations (Shahrokni, 2009; Yanguas, 2009; Yoshii & Flaitz, 2002). Second, pictures with word definitions support more efficient vocabulary learning than do other media combinations (Chun & Plass, 1996; Kost, Foss, & Lenzini, 1999; Yeh & Wang, 2003).

Captions for language learning

Captions in videos influence language learning. In general, use of video can "enrich target language processing" (Meskill, 1996, p. 196) and enhance listening ability (Brett, 1997), vocabulary acquisition (Brett, 1998; Duquette & Painchaud, 1996), and oral proficiency (Hung, 2009). Chapelle (2003) argued that modifications of the language and images in video content could increase learners' comprehension of the input. However, presentation of language and images alongside video content does not necessarily increase learners' comprehension (Danan, 2004). Research on multimedia use has focused on strategies for presentation of video materials (Baltova, 1994) rather than whether video represents an effective medium for learning. Specifically, researchers have investigated the effectiveness of on-screen text in the form of subtitles (L1), reversed subtitles (L2 text, L1 video), captioning (L2 video, L2 text), and keyword captioning (L2 video, L2 keywords). Overall, the results of these studies have indicated that all of the described forms are effective for vocabulary acquisition (Danan, 1992; Markham, 2001; Winke et al., 2010). In addition, studies have investigated the effects of captions and subtitles on learners' development of listening skills (Huang & Eskey, 2000; Markham & Peter, 2003) and reading comprehension (Garza, 1991; Goldman & Goldman, 1988).

Empirical studies have demonstrated that captions affect language learning. Hosogoshi (2016) investigated the effects of captions and subtitles on the listening process, specifically with regard to 11 listening strategies. In the study, a total of 114 Japanese-speaking EFL learners were divided into a no-text group, English caption group, and Japanese subtitle group. The findings indicated that the use of imagery and summarization strategies was significantly more prevalent in the Japanese subtitle group than other groups, and greater intercorrelations were observed between most of the listening strategies for each listening process in the no-text and caption groups than in the Japanese subtitle group. In a meta-analysis of the literature from a 30-year period, Montero Perez et al. (2013) reviewed more than 150 studies that employed captions. Their study revealed significant and large effect sizes for full captions (compared with no captions) in relation to video comprehension and vocabulary learning. Their findings also suggested that on-screen text can aid listening processes. In addition, Winke, Gass, and Sydorenko (2013) examined learners' caption

reading in various foreign languages and assessed the effect of familiarity with video content. Thirty-two participants with L1 English who were learning Arabic, Chinese, Russian, or Spanish watched two clips that presented familiar or unfamiliar content in their target language. The researchers determined that the learners looked at the captions for an average of 68% of the time that the captions were presented. Moreover, they determined that the various L2s were correlated with differences in learners' caption-reading behaviors. Learners of Arabic spent a higher proportion of their time looking at the captions compared with learners of Spanish or Russian.

The effect of subtitles and captions on vocabulary learning

Empirical studies have demonstrated that subtitles and captions can influence vocabulary learning. Regarding subtitles, Katchen (1997) conducted an empirical study to investigate potential instructional problems associated with the use of L2 television recordings with L1 subtitles in an L2 classroom. The findings revealed that the foreign language movies with subtitles benefitted learners' vocabulary. Additionally, Kosslstra and Beentjes (1999) examined the effectiveness of subtitles with L2 spoken videos for Dutch vocabulary acquisition and word recognition among fourth and sixth graders. Students in the experimental group who viewed the L2 spoken video with subtitles outperformed those in the control group in vocabulary learning.

Three empirical studies have emphasized the effect of captions on vocabulary learning. First, Bean and Wilson (1989) discovered that nonnative-speaker adult students held positive attitudes toward captioning, and their vocabularies increased when captioned materials were used. Students who viewed L2 captioned materials exhibited significant improvement in reading comprehension, listening comprehension, vocabulary acquisition, and word recognition. Second, Neuman (1990) designed four learning modes in an examination of the effectiveness of captioned materials. According to the study findings, participants who watched captioned programs learned more L2 words than did participants in the other three groups. Finally, Ellsworth (1996) conducted a study in which students were assigned to groups corresponding with three modes of presentation: (a) video with captions, (b) video (no captions), and (c) audio only. Grammar and vocabulary based on idioms, slang, and cultural expectations in the United States were incorporated into the lesson plans. The findings indicated that the captioned videos may have been beneficial for students who were insecure in their use of an L2.

The effectiveness of captions for L2 learners' vocabulary acquisition

Research on captioned video materials has often been focused on the effectiveness of captions for vocabulary acquisition among L2 learners. Preliminary studies have demonstrated that captioned video positively affects vocabulary learning (Bird & Williams, 2002; Chai & Erlam, 2008; Markham, 1999; Winke et al., 2010). Two general conclusions have been reported: captions significantly enhance (a) learners' recognition of written forms of vocabulary (Neuman & Koskinen, 1992; Sydorenko, 2010) and (b) their performance on aural form recognition tests (Markham, 1999). Montero Perez et al. (2014) compared learners' incidental vocabulary gains under four conditions, which differed in amount of on-screen text and the visual salience of keywords. Specifically, the groups received full captioning, keyword captioning, full captioning with highlighted keywords, or video only (control

group). The findings demonstrated that learners in the three groups that experienced captioned materials outperformed those in the control group in form recognition and, to a limited extent, meaning recognition of 17 target words.

Winke et al. (2010) reported that captions help L2 learners isolate and focus on word forms, and that consequently learners are more likely to notice these forms upon subsequent encounters. This finding is consistent with those of studies regarding the noticing of unknown words as the first step in the vocabulary acquisition process (Huckin & Coady, 1999; Hulstijn, 2001). Vanderplank (1990) presented a "speculative model" in which attention plays a crucial role in language learning through captioned video content (p. 228). In this model, the use of captions helps learners to carefully study the language spoken in a video by stimulating the "taking out" of language from the captioned video. This "taking out" of language occurs through both attention and adaptation. Attention is defined as a conscious selection process based on systematic "noting and gathering" of information and involves a reflective component in which learners notice a gap when comparing their L2 knowledge with captioned video input. By comparison, adaptation is defined as learner's selective attention to linguistic elements "for their own purposes" (p. 229).

Studies have used eye-tracking techniques to determine the amount of attention learners paid to video, audio, and captions. Eye-tracking can reveal how long learners spend looking at captions (Winke et al., 2013) and provide an objective measure for studying the role of attention in vocabulary learning (Godfroid, Housen, & Boers, 2010). Montero Perez et al. (2015) investigated the effects of two attention-enhancing techniques on L2 students' learning and processing of novel French vocabulary (target words) through video content with L2 captions. Eye-movement data and vocabulary tests were employed to evaluate the effects of caption type (full or keyword captioning) and test announcement. For measurement of the effects of test announcement, learners were either informed (intentional) or not informed (incidental) about vocabulary tests before test administration. The results revealed that learners in the keyword groups outperformed those in the other groups on the form recognition test, and analyses of learners' total fixation and second pass time on the target words revealed a significant interaction effect between caption type and test announcement. Additionally, the longer the participants focused on a given word, the more likely they were to correctly recognize it in subsequent encounters.

The present study

The purpose of the present study was to investigate the effect of captioned video with glosses on learners' incidental professional vocabulary acquisition from business-related materials. The present study aimed to answer the following research questions:

1. Do noncaptioned, captioned, and caption-gloss modes exhibit different effects on learners' retention of incidental professional vocabulary?
2. Does the caption-gloss mode result in differences in short- and long-term incidental professional vocabulary retention?
3. What are learners' attitudes toward use of the caption-gloss mode for incidental
4. professional vocabulary learning?

Method

Participants

Fifty sophomores from a level-one English (II) course at a business college participated in this study. Level one represented the highest among three levels of English courses, and students had been placed in this level based on their English scores on the university entrance exam. The average of their New Test of English for International Communication scores was 559, reflecting intermediate English ability. A total of 12 participants were from the Department of Business Administration, and 17 participants were from the Department of Accounting; the other 21 participants were from the Department of Wealth and Taxation Management. One-way repeated-measures analysis of variance (ANOVA) was adopted in this study based on its ability to control for factors that cause variations among subjects and thereby increase the power of the analysis. Because of this enhanced statistical power, the method offers a high probability of detecting an effect among relatively few subjects. Participants all signed consent forms indicating that they agreed to participate in this study.

Dependent and independent variables

The independent variables were the three caption modes: noncaptioned mode, captioned mode, and caption-gloss mode. The dependent variable was students' scores on measurements of business vocabulary, including immediate and 3-week-delayed posttests. Each posttest involved four test types: two production tests, comprising word spelling from Chinese to English (CtoE) and word spelling from English to Chinese (EtoC), and two recognition tests, comprising multiple-choice CtoE and multiple-choice EtoC questions. Nation (2001) stated that word knowledge includes knowledge of form, knowledge of meaning, and knowledge of use. In the present study, the four posttests were designed to examine participants' acquisition and retention of target word form and meaning. The participants' mother tongue was Chinese and their target language was English. Therefore, in addition to the English form and meaning of each target word, production and recognition tests were designed to examine participants' abilities to produce the Chinese equivalent of each target word, ensuring thorough assessment of the participants' acquisition and retention of the target words.

Materials

Video selection. The video content used in this study comprised three episodes of *Undercover Boss*, a reality television franchise series created by Stephen Lambert. The show depicts the experiences of senior executives working undercover in their own companies, investigating how their firms actually work, identifying areas in which the firms can be improved, and rewarding hard-working employees. Based on its business-related content, *Undercover Boss* is an appropriate show for students of business majors to watch to acquire English business vocabulary. The first video used in this study (Season 5 Episode 2) featured the CEO of Loehmann's, Steven Newman, working undercover in the lingerie section of one of his discount department stores. The second video (Season 5 Episode 12) was about Joe DiDomizio, president and CEO of Hudson Group, working undercover in his travel retail

company. The third video (Season 5 Episode 8) featured Travis Boersma, president and cofounder of Dutch Bros. Coffee, working undercover at an El Salvador coffee plantation from which his company bought coffee beans. Each video used in the study was a 15-minute clip that included eight target words. The average Flesch Reading Ease score of the three videos was determined to be 86.5, and their average Flesch–Kincaid grade level was calculated as 5.0. These scores correspond with conversational English for consumers and denote that the clips are easy to understand for nonnative English speakers. In this study, the researcher, who was also the teacher of the 50 participants, invited two additional English teachers with more than 10 years of experience in teaching students in business colleges to further examine the difficulty of the video clips for the English language learners. The three teachers determined the comprehension difficulty of the video clips based on the following factors: (a) all of the participants were nonnative speakers; (b) the participants' English abilities were intermediate; and (c) the three video clips featured conversational business English. The teachers concluded that the content of the video clips was of medium comprehension difficulty and was appropriate for intermediate-level students.

Target word selection. Target words were defined as words that were relevant to the understanding of each video clip, crucial to the meaning of a sentence or paragraph in each video clip, and with a definition related to business. Fifteen target words were selected from each video clip, and the three experienced English teachers identified the level of the selected words as approximate for upper-intermediate students. The teachers noted that some of the target words that were intermediate in their general meaning may have been particularly difficult for students to understand within a business context. For example, "associate" means connecting someone or something in your mind with someone or something else. However, within a business context, "associate" can also mean "business partner." Thus, the teachers classified some generally intermediate words as closer to upper-intermediate level in this study.

A pretest (vocabulary survey of the target words) was adopted to control for participants' prior knowledge of the target words. Specifically, the pretest was used to verify that the participants were not familiar with the target words. This test was administered to the participants 1 week before the experiment.

The pretest for each video clip contained 15 potential target words and 15 distracters (business words), the purpose of which was to lessen participants' impressions of the 15 potential target words. Based on the results of the target word pretest for each video clip, the eight most unfamiliar words of the original 15 were selected to be final target words. For the first video clip, the final target words were complacency (n.), iconic brand (n.), amenity (n.), irrelevant (adj.), associate (n.), author (v.), overwhelmingly (adv.), and merchandise (n.). For the second video clip, the final target words were legitimacy (n.), terminal (n.), penetrate (v.), constant (adj.), evolve (v.), seamless (adj.), incubator (n.), and renovate (n.). For the third video clip, the final target words selected were found (v.), franchise (n.), eligible (adj.), preserve (v.), contestant (n.), sacrifice (v.), declare (v.), and signature (n.). This study assessed learners' retention of these target words based on their productive and receptive knowledge. This approach followed Shahov's (2012) argument that knowledge of a word involves the ability to recognize it (receptive knowledge) and to use it correctly (productive knowledge).

Production procedure of the captioned video clip with glosses. Two students from the Department of Information Management helped in the five-step process through which

glosses were added to the target words in the captions for the caption-gloss mode. First, the captions were cut using Google2srt software and were saved as a caption-gloss file. Second, the target words were marked by the researcher in the newly created file. Third, the file was deposited into a media library using CyberLink PowerDirector Ultimate software. Fourth, captions were inserted into the caption-gloss video clip using the same software. Finally, each target word was circled in black and the corresponding glosses were inserted using the picture in the picture function. A sample of a captioned video clip with a gloss is presented in Figure 1.

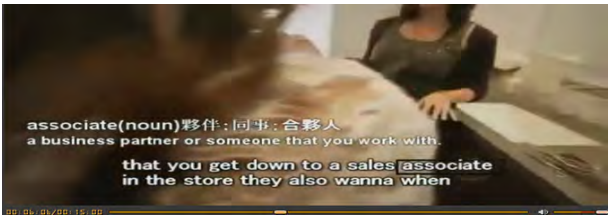

Figure 1. Sample of captioned video with a gloss

Instruments

Three video clip comprehension tests. For each video clip, one comprehension test was designed to trick the participants into believing they were being tested only on their comprehension, and not on their vocabulary knowledge. Each comprehension test consisted of five multiple-choice questions (two questions that assessed general understanding and three questions on content details).

A pretest. The pretest was a survey of the target vocabulary. All participants completed the pretest 1 week prior to the experiment. In their responses, the participants wrote down the Chinese equivalents of the English target words if they knew them; otherwise, they answered "I don't know." The purpose of the pretest was to confirm that all participants were unfamiliar with the target words.

Four vocabulary posttests. Immediate and 3-week-delayed posttests were conducted. Each posttest consisted of four types of vocabulary tests, and all eight target words appeared in each test (see Appendix). The first type of production test was CtoE word spelling. Participants were asked to spell the eight target English words based on the corresponding Chinese words, which were provided. The second type of production test was EtoC word spelling. Participants were asked to write the Chinese equivalents of the eight English target words, which were provided. For the recognition test based on multiple-choice CtoE questions, participants were provided with a Chinese word in each question and asked to choose the correct English equivalent from four options. Similarly, for the multiple-choice EtoC test, participants were provided with an English word in each question and asked to choose the corresponding Chinese word from four options. In the 3-week-delayed posttest, the order of the eight target words differed from the order in the immediate posttest to prevent participants from remembering the answers based on the sequence. For scoring, 1 point was awarded for each correctly spelled word in the word production tests, and 1

point was awarded for each correct answer in the word recognition tests. The highest score possible for each posttest was 8 and the lowest was 0.

Questionnaire. A questionnaire was designed to explore participants' attitudes toward the three caption modes (noncaptioned, captioned, and caption-gloss) used for incidental business vocabulary acquisition. The questionnaire consisted of two sections. The first section comprised three items that were ranked according to a 5-point Likert scale from 1 (strongly disagree) to 5 (strongly agree). The second section comprised three open-ended questions, each corresponding to one of the studied modes. The three items in the first section represented independent variables, and the results for these items were analyzed using descriptive statistics to determine the mode that was the most helpful for the participants' incidental professional word retention.

Treatment (modes). The study was conducted over a period of 6 weeks. As described, three distinct modes were tested. In mode one, participants used their cell phones to download a 15-minute noncaptioned video clip, and 30 minutes were allotted for them to watch the clip while answering five multiple-choice questions on a comprehension test. In mode two, participants downloaded a 15-minute captioned video clip using their cell phones, and 30 minutes were allotted for them to watch the clip while answering five multiple-choice questions on a comprehension test. In mode three, participants used their cell phones to download a 15-minute video clip with captions and glosses, and 30 minutes were allotted for them to watch the clip while completing five multiple-choice questions in a comprehension test.

Data collection procedure. One month before the experiment, each participant was provided with information about the study and signed a consent form agreeing to participate. In the first week of the study, participants completed a 30-minute pretest to verify their unfamiliarity with the target words in each video clip. The participants received over 20 minutes of training regarding how to download video clips from Google Cloud and how to use Google Docs to answer the comprehension-related questions. In the second week, participants completed the mode-one session within 30 minutes as well as a 10-minute posttest immediately thereafter. In the third week, the participants completed the mode-two session within 30 minutes immediately followed by another 10-minute posttest. In the fourth week, the participants completed the mode-three session within 30 minutes. They then completed an immediate posttest in 10 minutes and a 3-week-delayed posttest on mode-one words for another 10 minutes. In the fifth week, the participants were allotted 10 minutes to complete a 3-week-delayed posttest regarding mode-two words. Finally, in the sixth week, the participants completed a 3-week-delayed posttest regarding mode-three words within 10 minutes, and then they completed a questionnaire.

Data analyses. One-way repeated-measures ANOVA was used to analyze the collected data with regard to the null hypothesis and the first and second research questions. The independent variables were the three modes (noncaptioned, captioned, and caption glosses). The dependent variable was participants' scores on the business vocabulary posttests. The p value required for significance was .05 in all statistical analyses. To address the third research question, descriptive statistics and transcribed data were used to analyze the three 5-point Likert scale items and the three open-ended questions.

Results

Three modes and incidental professional vocabulary retention

The results of the pretests are presented in Table 1. For each video, fewer than 20% of the participants were familiar with any of the eight target words.

Table 1. Descriptive statistics of pretest by video type

Mode	*M*	*SD*
Noncaptioned	0.12	.44
Captioned	0.20	.50
Caption-gloss	0.30	.54

The scores on the immediate and 3-week-delayed posttests were regarded as the incidental target word retention outcomes. The immediate and 3-week-delayed posttests assessed how many target words the participants incidentally retained among the words presented in the three modes. As described, 1 point was credited for each correct answer, and the maximum score for each posttest was 8.

The descriptive statistics of the immediate posttest are presented in Table 2. In the immediate posttests, the participants scored highest for the target words displayed in the caption-gloss mode in both the CtoE production test (M = 1.55, SD = 1.33) and EtoC production test (M = 3.49, SD = 2.04). Moreover, the words corresponding to the caption-gloss mode were most frequently recognized in both the CtoE recognition test (M = 7.2, SD = 1.43) and EtoC recognition test (M = 7.57, SD = 1.06). Target word retention for words presented in the caption-gloss mode was the highest, followed by those for the captioned mode and the noncaptioned mode.

Table 2. Descriptive statistics of immediate posttest scores

	Production test				Recognition test			
	CtoE		EtoC		CtoE		EtoC	
Mode	*M*	*SD*	*M*	*SD*	*M*	*SD*	*M*	*SD*
Noncaptioned	0.08	0.34	0.39	0.8	5.14	2.06	6.78	1.47
Captioned	0.1	0.3	1.37	1.25	5.86	1.81	7.18	1.29
Caption-gloss	1.55	1.33	3.49	2.04	7.2	1.43	7.57	1.06

Note: CtoE, Chinese to English; EtoC, English to Chinese.

To determine significant differences between the effects of the three modes on learners' incidental vocabulary retention, one-way repeated-measures ANOVA was conducted with an alpha level of .05. The results of Mauchly's tests for the four posttests indicated that the assumption of sphericity was not met (p < .05). Therefore, the Greenhouse–Geisser correction was adopted.

Results revealed significant differences in participants' incidental vocabulary retention corresponding with the three modes in the CtoE production test (F(1.17, 58.71) = 53.80,

$p < .001$, $\eta^2 = 0.52$) and in the EtoC production test ($F(1.59, 79.66) = 59.49$, $p < .001$, $\eta^2 = 0.54$). To further examine the differences among the three modes in the two production tests, the pairwise comparison method was adopted. The results for the CtoE and EtoC production tests revealed that when the caption-gloss mode was employed, the participants retained significantly more of the target words than when the other two modes were employed. Additionally, the results of the EtoC production test revealed that the captioned mode resulted in significantly greater retention of target words than the noncaptioned mode.

A significant difference also existed among the three modes in results for the CtoE recognition test ($F(1.65, 82.45) = 17.27$, $p < .001$, $\eta^2 = 0.26$) and those for the EtoC recognition test ($F(1.76, 87.90) = 4.18$, $p < .05$, $\eta^2 = 0.08$). Results obtained using the pairwise comparison method revealed that participants' scores for the CtoE recognition test were highest to a significant degree when the caption-gloss mode was used. In terms of the EtoC recognition test, the caption-gloss mode helped participants to retain significantly more target words than the noncaptioned mode.

Table 3 lists the descriptive statistics of the 3-week-delayed posttests and the differences among the three modes, as determined using one-way repeated-measures ANOVA. In the 3-week-delayed posttests, participants scored the highest for the target words presented in the caption-gloss mode in both the CtoE production test ($M = 1.38$, $SD = 1.4$) and the EtoC production test ($M = 3.32$, $SD = 2.3$). Furthermore, target word retention for words from the caption-gloss mode was the highest in both the CtoE recognition test ($M = 6.88$, $SD = 1.7$) and the EtoC recognition test ($M = 7.5$, $SD = 1.4$). Participants acquired the most target words in the caption-gloss mode, followed by the noncaptioned mode and captioned mode.

Table 3. Descriptive statistics of the 3-week-delayed posttests

	Production test				**Recognition test**			
	CtoE		**EtoC**		**CtoE**		**EtoC**	
Mode	***M***	***SD***	***M***	***SD***	***M***	***SD***	***M***	***SD***
Noncaptioned	0.2	0.5	1.04	1.1	5.86	2.1	6.98	1.8
Captioned	0.12	0.4	1.72	1.6	5.86	1.9	6.7	2.1
Caption-gloss	1.38	1.4	3.32	2.3	6.88	1.7	7.5	1.4

Note: CtoE, Chinese to English; EtoC, English to Chinese.

Regarding the 3-week-delayed posttests, the assumption of sphericity was not met ($p < .001$) in the results of Mauchly's tests for the CtoE production test and the EtoC production test. Consequently, the Greenhouse–Geisser correction was used. Based on the result obtained using one-way repeated-measures ANOVA, the incidental vocabulary retention differed significantly between the three modes in the results for the CtoE production test ($F(1.22, 59.59) = 30.66$, $p < .001$, $\eta^2 = 0.39$) and those for the EtoC production test ($F(1.61, 79.06) = 20.62$, $p < .001$, $\eta^2 = 0.30$). To further investigate differences between the three modes, the pairwise comparison method was used. Results indicated that for both the CtoE and EtoC production tests, participants retained significantly more target words from the caption-gloss mode than from the other two modes.

Mauchly's tests indicated that the assumption of sphericity was met with $p = .57$ in the results for the CtoE recognition test and $p = .48$ in the results for the EtoC recognition test. Because these two Mauchly's test statistics for the CtoE and EtoC recognition test results

were not significant, the degree of freedom was interrupted directly without correction. A significant difference was observed between the scores for the three modes in the CtoE recognition test (F (2, 98) = 4.60, p < .01, η^2 = 0.09) and EtoC recognition test (F (2, 98) = 3.01, p < .05, η^2 = 0.06). Results obtained using the pairwise comparison method revealed that in the CtoE recognition test, participants retained significantly more incidental vocabulary when the caption-gloss mode was used than when either of the other two modes were used. In the EtoC recognition test, participants retained more incidental vocabulary when the caption-gloss mode was used than when the captioned mode was used.

In summary, the caption-gloss mode resulted in significantly higher performance than the other two modes in terms of learners' incidental vocabulary retention in three out of the four tests in the immediate and 3-week-delayed assessments. However, in the immediate EtoC recognition posttest, participants retained significantly more target words for the caption-gloss mode only when compared with results for the noncaptioned mode. No significant difference was identified in the mean scores for target word retention between the caption-gloss mode and captioned mode in the EtoC recognition posttest. Finally, the caption-gloss mode resulted in retention of significantly more target words than the captioned mode only in the 3-week-delayed EtoC recognition posttest. The mean scores for target word retention between the caption-gloss mode and noncaptioned mode in the EtoC recognition test were not significantly different.

The null hypothesis

The present study tested the null hypothesis – no significant relationship exists between vocabulary learning using captioned video with glosses and incidental professional vocabulary retention. Based on the analyses, the null hypothesis was rejected. The results in Tables 2 and 3 revealed a significant relationship between vocabulary learning in the caption-gloss mode and incidental professional vocabulary retention. The caption-gloss mode enabled learners to retain the most target words in the immediate and 3-week-delayed posttests and was the most beneficial for learners' incidental professional word acquisition among the three modes. These findings were supported by participant feedback. For example, one participant commented, *"The captions with glosses were good for learning new words and their usage."* Another participant stated, *"The glosses in the captions are a good way of enhancing my listening and vocabulary learning."* These findings were in accordance with expectations; several studies have determined that use of a multimedia glossary is an effective method for incidental vocabulary learning (Al-Seghayer, 2001; Hulstijn et al., 1996; Yanguas, 2009). Glosses facilitate vocabulary learning because they draw learners' attention to unknown words (Chun & Plass, 1996), and they enable learners to increase their vocabulary loads in small increments (Laufer & Hill, 2000).

Caption-glosses mode and incidental professional vocabulary retention in the short and long term

Regarding short-term results, the data presented in Table 2 revealed that participants retained the most target words from the caption-gloss mode. Regarding long-term results, the data presented in Table 3 indicated that use of the caption-gloss mode resulted in higher participant performance than the other two modes and led to retention of more target words. The results of paired sample t-tests revealed no significant difference in the

caption-gloss mode scores for the immediate production CtoE (M = 1.55, SD = 1.33) and the 3-week-delayed production CtoE (M = 1.45, SD = 1.47) posttests (t (49) = 0.36, p = 0.71). Additionally, no significant difference was observed in the caption-gloss mode scores for the immediate production EtoC (M = 3.49, SD = 2.04) and the 3-week-delayed production EtoC (M = 3.37, SD = 2.30) posttests (t (49) = 0.26, p = 0.79). Moreover, the results of paired sample t-tests indicated no significant difference in the caption-gloss mode scores for the immediate recognition CtoE (M = 7.20, SD = 1.42) and the 3-week-delayed recognition CtoE (M = 6.90, SD = 1.70) posttests (t (49) = 1.00, p = 0.32). Finally, no significant difference was identified in the caption-gloss mode scores for the immediate recognition EtoC (M = 7.57, SD = 1.06) and the 3-week-delayed recognition EtoC (M = 7.51, SD = 1.34) posttests (t (49) = 0.28, p = 0.77). These results suggested that participants' scores for incidental target word retention did not significantly decrease in the two 3-week production posttests and the two 3-week recognition posttests. Therefore, the caption-gloss mode may have enabled the participants to avoid significant loss of incidentally acquired target words over the 3-week period.

Learners' attitudes toward the caption-glosses mode in their incidental vocabulary learning

Following the completion of the 3-week-delayed posttest for the caption-gloss mode, participants were asked to use a 5-point Likert scale to score three questions regarding the helpfulness of the three modes for incidental professional target word retention. Table 4 lists the corresponding descriptive statistics, which revealed that the participants regarded the caption-gloss mode as the most helpful for incidental professional word retention (M = 4.02, SD = 0.89), followed by the captioned mode with a slightly lower mean of 3.86 (SD = 0.64), and the noncaptioned mode with the lowest mean of 2.92 (SD = 0.88).

Table 4. Descriptive statistics of learners' attitudes toward incidental professional word learning using the three modes

Item	*M*	*SD*
1. The noncaptioned mode was helpful for my incidental professional word retention.	2.92	0.88
2. The captioned mode was helpful for my incidental professional word retention.	3.86	0.64
3. The caption-gloss mode was helpful for my incidental professional word retention.	4.02	0.89

Table 5 presents the findings from participants' responses to the three open-ended questions regarding the most difficult and easiest parts of the learning process used as well as what they learned from the caption-gloss mode.

Regarding the most difficult part of the process, in spite of the glosses in the captions, some participants indicated that the speaker's Spanish accent (38%) and the unfamiliarity of many of the words used (26%) impeded their video comprehension. Regarding the easiest part, in addition to comprehension support from the images (18%), the most striking result was that most participants (92%) noted that the glosses in the captions facilitated their understanding of the video clip. Additionally, most participants (96%) reported that

they learned professional words from the glosses in the captions. In addition, the majority of participants (94%) indicated that the caption-gloss mode was a novel and relatively effective method for learning professional vocabulary and understanding video clips.

Table 5. Findings from three open-ended questions for the caption-gloss mode

Question items	Count	%
Q1: What was the most difficult aspect of watching the video?		
1. Some speakers' Spanish accents impeded my comprehension.	19	38%
2. The video featured too few familiar words to support my comprehension of the content.	13	26%
Q2: What was the easiest aspect of watching the video?		
1. The captions with glosses helped me understand the video.	46	92%
2. The images helped me understand the video.	9	18%
Q3: What did you learn by watching the video?		
1. I learned professional words from the glosses of the captions.	48	96%
2. Captions with glosses are a novel and effective method for learning professional vocabulary and understanding videos.	47	94%

Discussion

This study investigated the effect of the caption-gloss mode on learners' incidental professional word retention. The results suggested that (a) the caption-gloss mode resulted in retention of more incidental business words than did the noncaptioned and captioned modes; (b) among the three modes, the caption-gloss mode resulted in participant retention of the most target words in both the short- and long-term results; (c) no significant differences existed between the participants' short- and long-term retention when the caption-gloss mode was employed; and (d) of their attitudes toward the three modes, the learners exhibited the most positive attitude toward use of the caption-gloss mode for incidental business word acquisition.

Research question 1

The first research question assessed whether the noncaptioned, captioned, and caption-gloss modes would result in different effects on learners' retention of incidental professional vocabulary. The results indicated an affirmative answer. The data in Tables 2 and 3 indicate that the caption-gloss mode resulted in higher participant performance than did the other two modes in the immediate and 3-week-delayed posttests in terms of retention of incidental target words. Glosses in captions drew participants' attention to the target business words and enhanced their incidental acquisition of the words. This result was consistent with Hulstijn's (1992) claim that the provision of vocabulary glosses enhances incidental vocabulary learning. One student noted that they "*learned some professional business words from the glosses of the captions.*" Moreover, the result was supported by the finding of Montero Perez et al. (2018) that students in a glossed keyword caption group who were provided with meanings outperformed a no-captioning group, full-captioning group, and keyword-captioning group on form recognition and meaning recall tests. The results

of the present study were also in agreement with Aldera and Mohsen (2013), who reported that a group provided with animation, captions, and keyword annotations significantly outperformed a group provided with animation and captions on subsequent vocabulary recognition and vocabulary production tests. Similarly, in the present study, the glosses of the captions contributed to small increments in learners' vocabulary learning.

The method that the researcher used to mark the target words may have deepened participants' impressions of the target words. A participant indicated that "*the marked target words in the video made a deeper impression than other words.*" The findings of Montero Perez et al. (2014) supported the participants' reports that keyword captions or highlighted keywords may not only encourage learners to notice words, but also enhance learners' initial acquisition of word meaning.

Many participants preferred the caption-gloss mode because it combined video, audio, captions, and glosses. One participant noted, "*Glosses for professional words in the captions help me to understand the video content.*" Based on this type of feedback, glosses in captions were unlikely to distract learners and may have served as auxiliary aids that assisted participants in incidentally acquiring marked vocabulary. This inference is supported by the work of Jones (2003), whose study indicated that learners without access to annotations during a listening task were frustrated because they could not comprehend the input, whereas learners who were provided with annotations in various modalities did not express such concerns. Thus, a combination of inputs in multiple modalities may offer an advantage rather than a distraction for language learners.

Research question 2

The second research question concerned whether incidental professional vocabulary retention associated with the caption-gloss mode differs in the short and long term. Study results for the immediate and 3-week-delayed posttests indicated that participants retained the most target words when the caption-gloss mode was employed. According to the results of the paired samples test, no significant difference in target word retention existed between the immediate posttests and 3-week-delayed posttests for the caption-gloss mode. In the long term, participants' target word retention decreased slightly, but participants remembered most of the target words that they had acquired. Thus, short- and long-term results revealed that the caption-gloss mode was the most effective of the modes for assisting learners in incidental professional vocabulary acquisition. This finding was supported by the results of Abidin, Pour-Mohammadi, Shoar, Cheong, and Jafre (2011), who demonstrated that an electronic glossary in the form of multimedia annotation was more effective in the acquisition and learning of unknown vocabulary than non-electronic glossaries in forms such as printed text. Therefore, glosses in captions can be inferred to attract learners' attention in an incidental manner.

The difference in participants' target word retention between the immediate posttest and the 3-week-delayed posttest for the caption-gloss mode was slight, and results for the other two modes were similar. All target words across the three modes were related to the participants' business majors. Accordingly, the business-related meanings of words presented in glosses in the captions may have been relatively easy for participants to remember. In other words, because the meanings were related to the participants' majors, the target words might have left a deep impression, enhancing the participants' abilities to recall these words. This explanation echoes schemata theory. Many empirical studies have asserted

that schemata (background knowledge) are vital for EFL and ESL learners (Ajideh, 2003; Azizifar, Roshani, Gowhary, & Jamalinesari, 2015). Topical knowledge and familiarity with concepts behind unknown words have been identified as factors that influence incidental, context-based word learning (Forutan, 2011).

Research question 3

The third research question asked what attitudes participants held toward use of the caption-gloss mode for incidental professional vocabulary learning. As evident from the results presented in Tables 4 and 5, the participants regarded the caption-gloss mode as the most beneficial method of incidental professional vocabulary learning among the three modes. Participants' comments further supported this finding. One participant stated, "*I have learned the circled words with glosses in the captions.*" Another participant noted, "*A captioned video with glosses is a good method for acquiring new words.*" Furthermore, one participant stated that the glosses in the captions attracted her attention to the circled words: "*the circled words with glosses in the captions made a greater impression.*" These participant reactions were expected because studies (Al-Seghayer, 2001; Chun & Plass, 1996) have determined that multimedia annotations can draw participants' attention to target words. Therefore, glosses can be beneficial for incidental vocabulary learning.

Some participants indicated that although captions generally helped them understand the video content, the prevalence of unfamiliar or new words in the captions limited their understanding of the video. One participant noted, "*I do not have a large enough vocabulary, so I could not understand the video even with the captions. There were still some unknown words in the captions and they caused me to not understand some of the video content.*" Another participant made a similar comment: *"Even with the captions in the video, I still could not understand the new words because I did not understand some other unknown words apart from the new words.*" Additional data from the questionnaire verified that even with captions, participants were unable to fully comprehend the video content and vocabulary because other unknown words in the contextual clues hindered the students from learning the target words. Montero Perez, et al. (2018) indicated that the drawback of full captioning is that it does not explicitly offer the meaning of unknown words in the captions. Accordingly, in this study, the students relied on contextual clues to derive the meaning of unknown words. However, as described, additional unknown words obscured some of the context-based clues regarding the target words, inhibiting students' acquisition of the vocabulary. Most participants agreed that the glosses in the captions were crucial supports for incidental professional word acquisition and video comprehension. One participant stated, "*Glosses in the captions are a good way of enhancing video comprehension.*" Another participant noted, "*Looking at the glosses in the captions, I learned some professional business words and was able to understand the video better.*"

Limitations and suggestions for future studies

This study involved three primary limitations. First, because this study used a within-subjects design, carryover and fatigue effects were unavoidable. Second, stimulated recall should have been used to further verify the relationship between participant consultation of target words in the caption-gloss mode and their incidental acquisition of the words. Finally, participants used their cell phones to watch each video clip, and the small screens

may have resulted in participant fatigue during the 30-minute viewing periods. This possibility is supported by Stockwell and Liu (2015), who indicated that although the system in their study was designed to be read easily on smart phones, many learners felt that their phone screens were too small and the provided font size was not sufficiently large for them to properly complete the activities. Thus, if the present investigation is replicated in the future, an appropriate amount of time for watching video clips and suitable operational tools should be selected to avoid the fatigue problem. The benefits of the caption-gloss mode were demonstrated in the present study, implying potential directions for future studies. First, researchers might explore the effects of captions with glosses on incidental professional word retention in other fields such as law, engineering, and hospitality and tourism. Second, the mechanism by which the caption-gloss mode assists learners with incidental vocabulary acquisition in multimedia environments can be explored. Finally, the effect of the caption-gloss mode on learners' listening comprehension can be further elucidated.

Conclusion

This study investigated the effects of video content with captions and glosses on incidental professional vocabulary acquisition. Among the three tested modes, the caption-gloss mode resulted in participant retention of the most target words in the short and long term. Specifically, for this mode, participants' scores in the 3-week-delayed posttest were only slightly lower than those for the immediate posttest. Glosses in the captions attracted participants' attention and deepened their impressions of the target words. The target words were related to the participants' major, and this relationship may have also contributed to the participants' target word retention. Participants' positive feedback regarding the caption-gloss mode confirmed that the method was beneficial for their incidental vocabulary acquisition. The present study offers support for the use of captions with glosses in L2 incidental professional word learning. Teachers may use videos that include captions with glosses to draw students' attention to marked target words, thereby enhancing the students' retention of the vocabulary. Additionally, developers of professional English teaching materials can apply captions with glosses in teaching videos for incidental professional word learning.

In summary, of the three modes examined in this study, the caption-gloss mode was most effective for participant retention of target words. This result indicated that captions with glosses may attract L2 learners' attention in videos related to their professions, thereby enhancing incidental professional word learning. Based on these results, designers of ESP teaching material may consider integrating glosses into captions in teaching videos, and these findings may also serve as a reference for EFL teachers and learners.

References

AbdulGhani, S. (1993). ESP reading: Some implications for design of materials. *English Teaching Forum*, *7*, 42–45.

Abidin, M. J. Z., Pour-Mohammadi, M., Shoar, N. S., Cheong, S. T. H., & Jafre, A. M. (2011). A comparative study of using multimedia annotation and printed textual glossary in learning vocabulary. *International Journal of Learning and Development*, *1*(1), 82–90.

Ajideh, P. (2003). Schema theory-based pre-reading tasks: A neglected essential in the ESL reading class. *The Reading Matrix*, *3*(1), 1–14.

Aldera, A. S., & Mohsen, M. A. (2013). Annotations in captioned animation: Effects on vocabulary learning and listening skills. *Computers & Education*, *68*, 60–75.

Al-Seghayer, K. (2001). The effects of multimedia annotation modes on L2 vocabulary acquisition: A comparative study. *Language Learning and Technology*, *5*(1), 202–232.

Azizifar, A., Roshani, S., Gowhary, H., & Jamalinesari, A. (2015). The effect of pre-reading activities on the reading comprehension performance of Ilami High School Students. *Social and Behavioral Sciences*, *192*, 188–194.

Baltova, I. (1994). The impact of video on the comprehension skills of core French students. *The Canadian Modern Language Review*, *50*(3), 507–532.

Baltova, I. (1999). *The effect of subtitled and staged video input on the learning and retention of content and vocabulary in a second language*. (Unpublished doctoral dissertation). University of Toronto, Toronto, Canada.

Bean, R., & Wilson, R. (1989). Using closed-subtitled television to teach reading to adults. *Reading Research Instruction*, *28*(4), 27–37.

Bird, S. A., & Williams, J. N. (2002). The effect of bimodal input on implicit and explicit memory: An investigation into the benefits of within-language subtitling. *Applied Psycholinguistics*, *23*(4), 509–533.

Borras, I., & Lafayette, R. C. (1994). Effect of multimedia courseware subtitling on the speaking performance of college students of French. *The Modern Langauge Journal*, *78*, 61–75.

Brett, P. (1995). Multimedia for listening comprehension: The design of a multimedia-based resource for developing listening skills. *System*, *23*(1), 77–85.

Brett, P. (1997). A comparative study of the effects of the use of multimedia on listening comprehension. *System*, *25*(1), 39–53.

Brett, P. (1998). Using multimedia: A descriptive investigation of incidental language learning. *Computer Assisted Language Learning*, *11*(2), 179–200.

Burston, J. (2015). Twenty years of MALL project implementation: A meta-analysis of learning outcomes. *ReCALL*, *27*(1), 4–20.

Chai, J., & Erlam, R. (2008). The effect and the influence of the use of video and captions on second language learning. *New Zealand Studies in Applied Linguistics*, *14*(2), 25–44.

Chapelle, C. A. (2003). *English language learning and technology*. Philadelphia, PA: John Benjamins Publishing Company.

Chun, D. M., & Plass, J. L. (1996). Effects of multimedia annotations on vocabulary acquisition. *The Modern Language Journal*, *80*, 183–198.

Chung, T. M., & Nation, P. (2004). Identifying technical vocabulary. *System*, *32*, 251–263.

Danan, M. (1992). Reversed subtitling and Dual Coding Theory: New directions for foreign language instruction. *Language Learning*, *42*(4), 497–527.

Danan, M. (2004). Captioning and subtitling: Undervalued language learning strategies. *META*, *49*(1), 67–77.

Duquette, L., & Painchaud, G. (1996). A comparison of vocabulary acquisition in audio and video contexts. *The Canadian Modern Language Review*, *53*(1), 143–172.

Dwyer, M. (1978). *Strategies for improving visual learning*. State College, PA: Learning Services.

Ellsworth, T. (1996). Integrating subtitled video into your teaching. *English Teaching Forum*, *30* (3), 24-27.

Forutan, A. (2011). The impact of pre-reading and inserted questions on incidental vocabulary learning and retention. *Theory and Practice in Language Studies*, *1*(8), 980–985.

Fraser, S. (2006). The nature and role of specialized vocabulary: What do ESP teachers and learners need to know? *Hiroshima Studies in Language and Language Education*, *9*, 63–75.

Garza, T. J. (1991). Evaluating the use of captioned video materials in advanced foreign language learning. *Foreign Language Annals*, *24*(3), 239–258.

Gass, S. M. (1999). Discussion: Incidental vocabulary learning. *Studies in Second Language Acquisition*, *21*, 319–333.

Godfroid, A., Housen, A., & Boers, F. (2010). A procedure for testing the Noticing Hypothesis in the context of vocabulary acquisition. In M. Putz & L. Sicola (Ed.), *Cognitive processing in second language acquisition* (pp. 169–197). Amsterdam: John Benjamins.

Goldman, M., & Goldman, S. (1988). Reading with closed-captioned TV. *Journal of Reading*, *31*, 458–461.

Groot, P. J. M. (2000). Computer assisted second language vocabulary acquisition. *Language Learning and Technology*, *4*(1), 60–81.

Guillory, H. G. (1998). The effects of keyword captions to authentic French video on learner comprehension. *CALICO Journal*, *15*(1–3), 89–108.

Harding, K. (2007). *English for specific purposes*. Oxford: Oxford University Press.

Harji, M. B., Woods, P. C., & Alavi, Z. K. (2010). The effect of viewing subtitled videos on vocabulary learning. *Journal of College Teaching and Learning*, *7*(9), 37–41.

Hosogoshi, K. (2016). Effects of captions and subtitles on the listening process: Insight from EFL learners' listening strategies. *The JALT CALL Journal*, *12*(3), 153–178.

Huang, H., & Eskey, D. (2000). The effects of closed-captioned television on the listening comprehension of intermediate English as a second language students. *Educational Technology Systems*, *28*, 75–96.

Huckin, T., & Goady, J. (1999). Incidental vocabulary acquisition in a second language. *Studies in Second Language Acquisition*, *21*(2), 181–193.

Hulstijn, H., Hollander, M., & Greidanus, T. (1996). Incidental vocabulary learning by advanced foreign language students: The influence of marginal glosses, dictionary use, and reoccurrence of unknown words. *The Modern Language Journal*, *80*(3), 327–339.

Hulstijn, J. H. (1989). Implicit and incidental second language learning: Experiments in the processing of natural and partly artificial input. In H. W. Dechert & M. Raupach (Ed.), *Interlingual processes* (pp. 49–73). Tübingen, Germany: Gunter Narr.

Hulstijn, J. H. (1992). Retention of inferred and given word meanings: Experiments in incidental vocabulary learning. In P. J. L. Arnaud & H. Béjoint (Ed.), *Vocabulary and applied linguistics* (pp. 113–125). London: Macmillan.

Hulstijn, J. H. (2001). Intentional and incidental second language vocabulary learning: A reappraisal of elaboration, rehearsal and automaticity. In P. Robinson (Ed.), *Cognition and second language instruction* (pp. 258–286). Cambridge: Cambridge University Press.

Hung, H. T. (2009). Learners' perceived value of video as mediation in foreign language learning. *Journal of Educational Multimedia and Hypermedia*, *18*(2), 171–190.

Jacobs, G. M., Dufon, P., & Fong, C. H. (1994). L1 and L2 vocabulary glosses in L2 reading passages: Their effectiveness for increasing comprehension and vocabulary knowledge. *Journal of Research in Reading*, *17*(1), 19–28.

Jones, L. (2003). Supporting listening comprehension and vocabulary acquisition with multimedia annotations: The students' voice. *CALICO Journal*, *21*(1), 41–65.
Katchen, J. (1997). Can students learn English from the X-Files? In J. E. Katchen & N. N. Leung (Ed.), *The proceedings of the fifth of international symposium in English teaching* (pp. 243–250). Taipei: The Crane Publishing Company.
Ko, M. H. (2005). Glosses, comprehension, and strategy use. *Reading in a Foreign Language*, *17*(2), 125–143.
Kosslstra, C. M., & Beentjes., J. W. J. (1999). Children's vocabulary acquisition in a foreign language through watching subtitled television programs at home. *Educational Technology Research and Development*, *47*(1), 51–60.
Kost, C. R., Foss, P., & Lenzini, J. J. (1999). Textual and pictorial glosses: Effectiveness on incidental vocabulary growth when reading in a foreign language. *Foreign Language Annals*, *32*, 89–97.
Laufer, B. (2010). Form focused instruction in second language vocabulary learning. In R. Chacon-Beltran, C. Abello-Contessa, M. M. Torreblanca-Lopez, & M. D. Lopez-Jimenez (Ed.), *Further insights into non-native vocabulary teaching and learning* (pp. 15–27). Clevdeon, UK: Multilingual Matters.
Laufer, B., & Hill, M. (2000). What lexical information do L2 learners select in a CALL dictionary and how does it affect word retention? *Language Learning and Technology*, *3*(2), 58–76.
Magasic, M. (2017). Learning through watching: Streaming video in L2 English. *The JALT CALL Journal*, *13*(3), 199–209.
Markham, P. (1999). Captioned videotapes and second-language listening word recognition. *Foreign Language Annals*, *32*(3), 321–328.
Markham, P. (2001). The influence of culture-specific background knowledge and captions on second language comprehension. *Journal of Educational Technology Systems*, *29*(4), 331–343.
Markham, P., & Peter, L. (2003). The influence of English language and Spanish language captions on foreign language listening/reading comprehension. *Journal of Educational Technology Systems*, *31*(3), 331–341.
Marzban, A. (2011). Investigating the role of multimedia annotations in EFL reading comprehension. *Procedia-Social and Behavioral Sciences*, *28*, 72–77.
Meskill, C. (1996). Listening skills development through multimedia. *Journal of Educational Multimedia and Hypermedia*, *5*(2), 179–201.
Montero Perez, M., Noortgate, W. V. D., & Desmet, P. (2013). Captioned video for L2 listening and vocabulary learning: A meta-analysis. *System*, *41*(3), 720–739.
Montero Perez, M., Peters, E., Clarebout, G., & Desmet, P. (2014). Effects of captioning on video comprehension and incidental vocabulary learning. *Language Learning & Technology*, *18*, 118–141.
Montero Perez, M., Peters, E., & Desmet, P. (2015). Enhancing vocabulary learning through captioned video: An eye-tracking study. *The Modern Language Journal*, *99*(2), 308–328.
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.
Nation, I. S. P. (1999). *Learning vocabulary in another language*. New Zealand: Victoria University of Wellington.

Nation, I. S. P. (2001). *Learning vocabulary in another language*. Cambridge: Cambridge University Press.

Neuman, S. (1990). *Using subtitled television to improve the reading proficiency of language minority students*. Falls Church, VA: National Captioning Institute.

Neuman, S. B., & Koskinen, P. (1992). Captioned television as comprehensible input: Effects of incidental word learning from context for language minority students. *Reading Research Quarterly, 27*(1), 95–106.

Paivio, A. (1986). *Mental representations: A dual-coding approach*. New York, NY: Oxford University Press.

Paivio, A. (1991). Dual coding theory: Retrospect and current status. *Canadian Journal of Psychology, 45*, 255–287.

Paivio, A. (2007). *Mind and its evolution: A dual coding theoretical approach*. Mahwah, NJ: Erlbaum.

Peters, E., & Montero Perez, M. (2015). Vocabulary acquisition through audio (-visual) input: How to measure learning gains? Paper presented at the workshop on reliability and validity in SLA research, Barcelona.

Plass, J. L., Chun, D. M., Mayer, R. E., & Leutner, D. (1998). Supporting visual and verbal learning preferences in a second-langauge multimedia learning environment. *Journal of Educational Psychology, 90*, 25–36.

Sanders, C. S. (2002). *Integrating multimedia in technology education to improve college student comprehension, problem-solving skills, and attitudes toward instructional effectiveness*. (Unpublished doctoral dissertation). North Carolina State University, North Carolina.

Schmidt, R. (1994). Deconstructing consciousness in search of useful definitions for applied linguistics. *AILA Review, 11*, 11–26.

Shahrokni, S. A. (2009). Second language incidental vocabulary learning: The effect of online textual, pictorial, and textual pictorial glosses. *Teaching English as a Second or Foreign Language, 13*(3), 1–17. Retrieved from http://www.tesl-ej.org/wordpress/issues/volume13/ej51/ej51a3

Shahov, V. P. (2012). Measuring L2 receptive and productive vocabulary knowledge. *Language Studies Working papers, 4*, 37–45.

Skehan, P. (1998). *A cognitive approach to language learning*. Oxford: Oxford University Press.

Stockwell, G., & Liu, Y. C. (2015). Engaging in mobile phone-based activities for learning vocabulary: An investigation in Japan and Taiwan. *CALICO Journal, 32*(2), 299–322.

Sydorenko, T. (2010). Modality of input and vocabulary acquisition. *Language Learning & Technology, 14*(2), 50–73. Retrieved from http://llt.msu.edu/voll4num2/sydorenko.pdf

Vanderplank, R. (1990). Paying attention to the words: Practical and theoretical problems in watching television programs with unilingual (CEEFAX) sub-titles. *System, 18*(2), 221–234.

Vanderplank, R. (2010). A decade of research on language laboratories, television and video in language learning. *Language Teaching, 43*(1), 1–37.

Vanderplank, R. (2013). Effects of and "effect with" captions: How exactly does watching a TV program with same-language subtitles make a difference to language learners? *Language Teaching, Available on CJO 2013*, 1–16.

Watanabe, Y. (1997). Input, intake, and retention: Effects of increased processing on incidental learning of foreign language vocabulary. *Studies in Second Language Acquisition*, *19*, 287–307.
Winke, P., Gass, S., & Sydorenko, T. (2010). The effects of captioning videos used for foreign language listening activities. *Language Learning & Technology*, *14*(1), 65–86.
Winke, P., Gass, S., & Sydorenko, T. (2013). Factors influencing the use of captions by foreign language learners: An eye-tracking study. *The Modern Language Journal*, *97*(1), 254–275.
Yanguas, I. (2009). Multimedia glosses and their effect on L2 text comprehension and vocabulary learning. *Language Learning and Technology*, *13*(2), 48–67.
Yeh, Y., & Wang, C. (2003). Effects of multimedia vocabulary annotations and learning styles on vocabulary learning. *CALICO Journal*, *21*, 131–144.
Yoshii, M. (2006). L1 and L2 glosses: Their effects on incidental vocabulary learning. *Langauge Learning & Technology*, *10*(3), 85–101.
Yoshii, M., & Flaitz, J. (2002). Second language incidental vocabulary retention: The effect of text and picture annotation types. *CALICO Journal, 20*, 33–58.

Appendix A

Sample of the four posttests

(A.1) Production test (CtoE)
Please write down the equivalent English word.
請按照中文意思，寫出所屬的英文單字。
___ 1. 便利設施;福利設施
___ 2. 知名的;受歡迎的
___ 3. 滿足;自信
___ 4. 開創;發起
___ 5. 壓倒性地;不可抵抗地
___ 6. 無相關的;不恰當的
___ 7. 夥伴;同事;合夥人
___ 8. 商品;貨物

(A. 2) Production test (EtoC)
Please write down the Chinese equivalent.
請按照英文單字，寫出所屬的中文意思。
___ 1. irrelevant
___ 2. iconic brand
___ 3. overwhelmingly
___ 4. amenity
___ 5. complacency
___ 6. associate
___ 7. author
___ 8. merchandise

(A. 3) Recognition test (CtoE)
Please select the correct English equivalent for the given Chinese.

請按照中文意思，選出正確的英文單字。

1. 夥伴; 同事; 合夥人
   (A) associate (B) tension (C) specialist (D) statement
2. 開創;發起
   (A) inhale (B) persist (C) implement (D) author
3. 便利設施; 福利設施
   (A) customer (B) bankruptcy (C) e-commerce (D) amenity
4. 壓倒性地; 不可抵抗地
   (A) exactly (B) functionally (C) overwhelmingly (D) transitionally
5. 滿足; 自信
   (A) reference (B) complacency (C) expense (D) entrepreneur
6. 知名的; 受歡迎的
   (A) front line (B) brand fashion (C) iconic brand (D) thought-provoking
7. 無相關的;不恰當的
   (A) irrelevant (B) major (C) premium (C) aspiring
8. 商品; 貨物
   (A) differentiator (B) performance (C) merchandise (D) laments

(A. 4) Recognition test (EtoC)

Please select the correct Chinese equivalent for the given English word.

請按照英文單字，選出正確的中文意思。

1. irrelevant
   (A) 熱情的; 積極的 (B) 無相關的; 不恰當的 (C) 能力的;實力的 (D) 著重的; 專注的
2. amenity
   (A) 便利設施;福利設施 (B) 預估; 估計 (C) 參考; 推薦人 (D) 彈力; 彈性
3. overwhelmingly
   (A) 實際地; 事實上 (B) 壓倒性地;不可抵抗地 (C) 絕對地; 完全地 (D) 正確地; 正好地
4. author
   (A) 執行; 進行 (B) 監控; 監視 (C) 參與; 加入 (D) 開創; 發起
5. iconic brand
   (A) 時尚的; 時髦的 (B) 靈機一動的 (C) 獨特的; 特別的 (D) 知名的; 受歡迎的
6. associate
   (A) 夥伴; 同事 (B) 利益; 好處 (C) 破產;倒閉 (D) 區分者; 微分器
7. merchandise
   (A) 商品; 貨物 (B) 執行長; 總裁 (C) 監督者; 管理者 (D) 挽歌; 悼詞
8. complacency
   (A) 集團; 企業群 (B) 滿足;自信 (C) 費用; 花費 (D) 經營; 運作