Incidental vocabulary acquisition from documentary viewing: A study on the role of captions[1]

Adquisición incidental de vocabulario a partir de la visualización documental: un estudio sobre el papel de los subtítulos

Marion Durbahn [2]

Abstract

This study investigated the incidental acquisition of second language (L2) vocabulary through the viewing of a 30-minute science documentary and the effect of captioning on this process. Thirty-four English language learners watched two documentaries (one captioned and one uncaptioned) and were then assessed on their knowledge of a set of words appearing in the documentaries. Results of the Vocabulary Tests (Form-recall/Spelling, and Meaning recognition) showed that documentary viewing led to significant acquisition of new vocabulary and that the captioning condition had a significant effect on this learning. Captioning boosted the gains on meaning recognition, form recall and spelling. Results indicated that relative gains in the captioning condition are 8% significantly higher than in the uncaptioned condition.

Keywords: vocabulary, incidental learning, captioning, recall, recognition.

Resumen

Este estudio investigó la adquisición incidental de vocabulario de segunda lengua (L2) mediante la visualización de un documental científico de 30 minutos y el efecto de los subtítulos en este proceso. Treinta y cuatro aprendices del idioma inglés vieron dos documentales (uno con subtítulos y otro no subtitulado) y luego se les evaluó el conocimiento de un conjunto de palabras que aparecen en los documentales. Los resultados de las Pruebas de vocabulario (reconocimiento de formas / deletreo y reconocimiento de significado) mostraron que la visualización documental llevó a una adquisición significativa de vocabulario nuevo y que la condición de subtitulado tuvo un efecto significativo en este aprendizaje. Los subtítulos aumentaron el reconocimiento del significado y posibilitaron la recuperación de formas y la ortografía. Los resultados indicaron que las ganancias relativas en la condición de subtitulado son un 8% significativamente más altas que en la condición no subtitulada.

Palabras clave: vocabulario, aprendizaje incidental, subtitulación, recuerdo, reconocimiento.

---

[1] Some ideas in this paper come from Marion Durbahn's MA thesis in Applied Linguistics with ELT, submitted to the University of Nottingham, UK.

[2] Katholieke Universiteit Leuven, Belgium. Email: marion.durbahn@kuleuven.be

Introduction

The advent of the communicative approach has naturally led to a focus on incidental learning (DeCarrico, 2001). Since then, it has become a particularly productive area of second language (L2) vocabulary acquisition research. This kind of learning has been described as a 'side effect' (Gass, 1999) of any communicative activity not explicitly geared towards vocabulary learning, or as a 'by-product' of the main cognitive activity (Huckin & Coady, 1999). Overall, studies have shown that new vocabulary can be learnt incidentally from reading (Brown, Waring, & Donkaewbua, 2008; Pellicer-Sánchez & Schmitt, 2010; Pigada & Schmitt, 2006; Rott, 1999; Vidal, 2011; Waring & Takaki, 2003; Webb, 2007), listening (Brown, et al., 2008; van Zeeland & Schmitt, 2013; Vidal, 2003, 2011), reading-while-listening (Brown, et al., 2008; Neuman & Koskinen, 1992, and Webb & Chang, 2015) and video viewing (Neuman & Koskinen, 1992; Huang & Eskey, 1999; Markham, 1999; Webb & Rodgers, 2009 a, 2009b; Rodgers & Webb 2011; Rodgers, 2013; Montero Perez, Peters, Clarebout, & Desmet, 2014; Peters, Heynen, and Puimège, 2016; Peters & Webb, 2018).

Nation (2007) posits that one of the main conditions for incidental learning to take place is that learners must receive large amounts of input that are comprehensible and enjoyable for them. This is to draw learners' attention for longer periods and provide repeated encounters (Day & Bamford, 1998). However, programs designed to promote incidental L2 vocabulary acquisition often employ materials adapted for language learners, which do not often include vocabulary beyond the 3,000 or 4,000 most frequent words (Rodgers, 2013). This has led teachers and researchers to resort to more authentic material to foster vocabulary acquisition of lower frequency words (Schmitt & Schmitt, 2014). That is why the use of authentic audio-visual material for incidental learning has recently attracted the attention of vocabulary researchers. The advantage of audio-visual input over other input modes lies in the fact that audio is supported by imagery, and, on some occasions, different forms of on-screen text are present. This combination of verbal and visual input is believed to support the comprehension process and improve learning gains (Mayer, 2009). Overall, studies have shown the effectiveness of audio-visual materials for the acquisition of new words at different levels of word knowledge (Nation, 2001), i.e., meaning recognition, form recognition, meaning recall, and form recall (Hui, 2007; Markham, 1999; Markham, Peter & McCarthy, 2001; Rodgers, 2013; Montero Perez, Peters, & Desmet, 2015; Montero Perez, et al., 2014; Peters, Heynen, & Puimège, 2016; Peters & Webb, 2018)

Previous studies have also shown that the incidental learning from audio-visual materials can be further supported by the use of captions, and that they significantly aid word form recognition (Neuman & Koskinen, 1992; Sydorenko, 2010). However, the effect of captions (L2 audio plus L2 text) on other aspects of lexical knowledge, such as meaning recognition and form recall, is still largely unexplored. Importantly, with the very few studies available, the effect of captions on the incidental vocabulary learning from different types of materials and in different contexts is still unclear.

Broadly, this study comprises a brief description of the state of the art on incidental vocabulary learning and viewing comprehension together with conceptual definitions of the variables under investigation. It continues with the research questions and the methodology employed to carry out the study, which includes a description of the participants, the selection of the materials and target words, the instruments design, and the procedure. Next, the analysis and discussion of the results are presented. To finalize, the conclusion provides a summary of the main findings along with the shortcomings and limitations of the study design.

Background

Vocabulary knowledge

There is no doubt that vocabulary knowledge is an important component of language use and an essential part of the second language learning mastering. It has been found to be a great predictor of language performance, overall in reading comprehension, where it has been found that the larger the learners' vocabulary size, the higher the levels of comprehension. However, vocabulary size (also called vocabulary breath) is not enough to independently function in a second or foreign language (Schmitt, 2010). Learners must also know how each individual item behaves when they are together with other items. This kind of knowledge is often referred to as vocabulary depth, which is a rich and complex construct with different dimensions. Nation (2001) conceptualized vocabulary depth by breaking it down in its different components, as shown in Figure 1.

Figure 1. *Aspects involved in knowing a word*

Form	spoken	R	What does the word sound like?
		P	How is the word pronounced?
	Written	R	What does the word look like?
		P	How is the word written and spelled?
	word parts	R	What parts are recognizable in this word?
		P	What word parts are needed to express the meaning?
Meaning	form and meaning	R	What meaning does this word form signal?
		P	What word form can be used to express this meaning?
	concept and reference	R	What is included in the concept?
		P	What items can the concept refer to?
	associations	R	What other words does this make us think of?
		P	What other words could we use instead of this one?
Use	grammatical functions	R	In what patterns does the word occur?
		P	In what patterns must we use this word?
	collocations	R	What words or types of words occur with this one?
		P	What words or types of words must we use with this one?
	constraints on use	R	Where, when, and how often would we expect to meet this word?
	(register, frequency…)	P	Where, when, and how often can we use this word?

Note: R = receptive, P = productive. Adapted from Nation (2001, p. 27)

Schmitt (2010) claims that the form-meaning link is relatively easy to measure for teaching and researching purposes whereas other aspects, such as collocations or register are extremely difficult. The form-meaning connection dimension is, in turn, subdivided into receptive and productive. While receptive vocabulary can be measured as form or meaning recognition, productive vocabulary can be measured as form and meaning recall. In recognition tasks, the learners have to select the correct meaning or the correct form of a word among different alternatives whereas in recall tasks, learners have to produce the meaning or the form of a word (Schmitt, 2010). Meaning and form recognition and recall are the most common components assessed when dealing with incidental vocabulary learning, at least for single words.

Incidental vocabulary learning

Language acquisition can occur either intentionally or incidentally. Intentional and incidental learning are two mainstream concepts in research not only on second language (L2) acquisition, but also on L2 pedagogy (Bruton, López, & Mesa, 2011). Intentional learning is the result of a deliberate effort to study the language to commit information to long-term memory (Hulstijn, 2001; Laufer & Hulstijn, 2001), where rehearsal and memorizing techniques play a fundamental role (Schmitt, 1997). Incidental learning, on the other hand, takes place when learners learn new features of a language without intentional exposure to them (Hulstijn, 2003). In other words, learning occurs naturally (Bruton et al., 2011) while the focus is on the development of other abilities, usually comprehension (Rodgers, 2013), or even while engaging in communicative activities for pleasure, such as watching television, reading magazines, or listening to music.

Research on incidental vocabulary learning has traditionally focused on the acquisition of new words from reading (with a few studies looking at reading-while-listening and listening). However, generalizing the results from studies on reading and listening seems not to be appropriate, overall, because audio-visual input contains imagery, which provides support to grasp the meaning of unknown words and, thus, aid incidental vocabulary learning. To date, there are some corpus-based studies (Rodgers & Webb, 2011; Webb & Rodgers, 2009a, 2009b; Webb, 2011) that have shown that it may be possible to acquire words incidentally from multimedia material. Empirical studies investigating the benefits of video viewing on incidental vocabulary acquisition have by and large explored the effect of captioned videos (audio supported by text) and have looked at learning gains in different dimensions of vocabulary knowledge.

Incidental vocabulary learning from captioned videos

Multimedia materials offer a wide variety of support to the audio that usually appears simultaneously in the form of text at the bottom of a television or screen (Chung, 1999). The name and purpose of the support vary according to the language in which they are presented, such as audio only (either L1 or L2), subtitles (L2 audio, L1 on-screen text), reversed subtitles (L1 audio, L2 text), and captions(L2 video, L2 text) (Montero Perez, et al., 2014). Among these different conditions, captions, also referred to as bimodal, unilingual and intralingual subtitles (Danan, 2004), have particularly received the attention from vocabulary researchers because of their potential to support the learning process. Captions are acknowledged for enhancing vocabulary learning in a similar way, or even more efficiently, than reading (Rodgers, 2013), listening (Baltova, 1999, and Markham, 1999), and reading-while -listening (Neuman & Koskinen, 1992) since learners usually rely on captions when they do not know or understand the meaning or pronunciation of unknown words (Rodgers, 2013). Captions aid speech decoding and segmentation, facilitating the recognition of word boundaries (Montero Perez et al., 2014). They can also help learners cope with input that is slightly higher than their level of proficiency because they facilitate the form-meaning mapping (Danan, 2004). While the bulk of studies have investigated the influence of captions on video comprehension (e.g. Baltova, 1999; Chung, 1999; Huang & Eskey, 1999; Neuman & Koskinen, 1992; Winke, Gass, and Sydorenko, 2013), fewer studies have examined their effect on incidental vocabulary learning.

Huang and Eskey (1999) examined the effect of captioned over uncaptioned videos on incidental vocabulary learning. Thirty intermediate ESL students viewed a 14 -minute television series designed for ESL classroom teaching and were then assessed on their knowledge of 16

target words. The results of the study showed that the captioned group scored significantly higher than the uncaptioned group on the vocabulary tests. Similar results were found by Markham (1999), where 118 advanced ESL university students viewed two excerpts from educational television programs and were tested on their knowledge of 100 target words by means of a form recognition multiple-choice test. Results indicated that the availability of captions significantly aided the participants' ability to recognize the target word forms. Likewise, Neuman and Koskinen (1992) examined the effect of captioned videos on the incidental acquisition of form recognition, meaning recognition, and form recall. A sample of 179 seventh and eighth graders viewed 9 television segments and were assessed on their knowledge of 90 target words. Findings revealed that the captioning group significantly outperformed the uncaptioned group on meaning recognition (see Appendix A for a summary of these studies).

The effect of captions has also been compared to other modes of support in audio-visual materials. For example, Baltova (1999) compared the effect of captions and subtitles on vocabulary acquisition. In her study, participants were asked to view a brief video documentary in one of three viewing conditions: Bimodal (captions), Reversed format (subtitles), and Traditional format (No Captions). Knowledge of thirty target words were measured by means of a Video C - Cloze test and the Vocabulary Knowledge Scale. Results indicated that the vocabulary gains and retention rates were found to be significantly higher in the Bimodal than in the other conditions. The same viewing conditions were examined in Hui (2007), where major students of English viewed a science documentary and were assessed on their knowledge of form recognition, meaning recall, and spelling. Findings showed that captions significantly fostered participants' word spelling and oral word- recognition. Also, captions and subtitles appeared to be more beneficial for both high and low L2 proficiency students than audio only. Regarding meaning recognition, Hui found that low proficiency learners benefit more from subtitles than from captioning. Thus, the author advises caution when introducing captions in the language classroom. A more recent study conducted by Peters, Heynen, and Puimège (2016) concluded that captions lead to higher gains than subtitles on form recognition. However, no difference between captions and subtitles were found regarding word meaning recall.

The effect of different types of captions has also been explored. Montero Perez, Peters, and Desmet (2015) and Montero Perez, Peters, Clarebout, and Desmet (2014) looked at the keyword method, an attention-enhancing technique which consists of a visual salience in the captioning line, where the target word appears in isolation. In the first study, 51 Dutch-speaking undergraduate students at a Flemish university watched two authentic French clips from a Swiss and Belgian current affairs program for native speakers of French. They studied the effect of captioning and the keyword method on incidental and intentional vocabulary learning using a combination of eye-movement data and four vocabulary tests aiming at measuring form recognition, meaning recognition, meaning recall, and clip association. Results showed that learners in the keyword groups significantly outperformed the other groups on the form recognition test. In the second study, 133 Flemish undergraduate students watched three French clips. A fourth condition was included in the design: Fully captioned clips with highlighted words. After the viewing session, the participants completed four vocabulary tests measuring form recognition, meaning recognition, meaning recall, and clip association. Results showed that both keyword groups outperformed the fully captioning and no captioning groups. A summary of the findings described above is provided in Appendix A.

Overall, the studies reviewed have shown that new vocabulary can be learnt incidentally from different types of video viewing and that captions seem to support this process and further enhance the learning potential of audio-visual materials. Importantly, the use of captions seems to have an advantage over other techniques such as subtitles (although this effect seems to be modulated by learners' proficiency). However, a number of important methodological concerns need to be considered. Regarding the type of word knowledge aspects assessed, most of the studies reviewed (see Appendix B) looked at word form recognition and gains range from 23% to 69.5%, with a mean close to 50%. A few of them report gains on meaning recognition (3.5% to 61.4%) and meaning recall (.82% to 46.7%). Only two of them examined form recall and just one looks at spelling. Learning gains reported in the studies have shown a huge variation, making the effect of captioned videos for the incidental learning of different aspects of vocabulary knowledge still unclear.

Another important methodological issue concerns the type of stimuli used in these studies. These studies used episodes of a television series, educational programs, current affairs programs, television segments for children, episodes from a comedy series, and video-taped lectures. Despite the benefits attributed to authentic materials for incidental learning, few studies have employed documentaries as stimuli (Baltova, 1999; Markham, et al., 2001; Hui, 2007, and Peters, et al., 2016), but their viewing time is fairly short ranging from 7 to 20 minutes. The short viewing time is a major weakness which affects the implications for formal and informal language learning (Vanderplank, 2010). Coincidently, gains from studies with longer viewing time looking at incidental vocabulary learning, such as Rodgers (2013, 8 hours), are usually smaller, which points to the need of more studies on incidental vocabulary learning from viewing with longer viewing time.

In order to address these gaps, the present study examined the extent to which captioning affects the incidental acquisition of different aspects of lexical knowledge, namely word form recall/spelling and meaning recognition. Accordingly, the central questions under investigation are:

1. How much vocabulary are participants able to learn incidentally from uncaptioned and captioned video viewing, as measured by Form-recall/Spelling and Meaning-recognition Tests?
2. Does captioning lead to higher incidental L2 vocabulary gains than words presented in audio-visual input with no captions, as measured by Form-recall/Spelling and Meaning-recognition Tests?

Methodology

Participants

Thirty-eight English language learners with Spanish as their first language (L1) initially participated in the study. Four participants were excluded from the analysis because of either Spanish not being their L1 or they did not complete the experimental procedure. The resulting number was 34 participants (12 males and 22 females), whose ages ranged from 21 to 37 years old ($M = 27.7$; $SD = 4.2$). They were all learners of English and had taken an advanced English course. They were studying English in two different contexts (16 EFL and 18 ESL).

All participants completed a Yes/No vocabulary size test (Meara, n.d.) to assess their vocabulary size and account for potential differences in their proficiency. Participants had an average vocabulary size of 6.139 (*SD* = 1.015.5, *Min.* = 4.255, *Max.* = 7.929) word families. An independent-samples t-test showed that there were no significant differences between the scores of participants in the EFL (*M* = 6.219; *SD* = 931) and ESL (*M* = 6.075; *SD* = 1.098) contexts ($t(32)$= .41, *p* = .69).

Materials

Video selection

The videos selected for the study were two episodes from the NOVA's documentary series: *Making Stuff* and *Making More Stuff* (NOVA, 2011, 2013). It is a series of two seasons of four episodes each that discusses scientific innovations in materials science using language that any person, not necessarily familiar with the topic, could easily understand. All participants viewed the two documentaries. The first documentary, *Making Stuff Wilder* (Doc 1, season 2, uncaptioned), explores bold technological innovations inspired by nature. The second, *Making Stuff Smarter* (Doc 2, season 1, captioned), deals with engineering materials that can shape themselves. Each of the documentaries was assigned to one of the two experimental conditions (uncaptioned or captioned). The original documentaries lasted 54'25'' and 53'24'', respectively. However, because of practical reasons and the fact that not all the topics contained in the videos were useful for the purpose of this study, their lengths were reduced to 32'03'' (4367 tokens, 1193 types, 873 families) and 31'38'' (4166 tokens, 1178 types), respectively. The documentaries were chosen because they contained technical words that may be unknown to the participants.

A lexical frequency analysis, using Cobb's (n.d.) Lextutor website, revealed that, if proper nouns and interjections are excluded, both documentaries had a similar lexical density (Table 1).

Table 1. *Lexical frequency profile in Doc 1 and Doc 2*

Lexical Frequency Profile						
	1K*	2K	3K	4K	5K	10K
Doc 1	77.6%	86.3%	92.6%	95%	95.9%	97.8%
Doc 2	78.5%	88.2%	92.5%	94.2%	95.6%	97%

*K representing 1,000 word families

Nation (2006), Laufer and Ravenhorst-Kalovski (2010) and van Zeeland and Schmitt (2013) suggest that comprehension can take place when learners understand at least 95% of a text (although 98% is better). Table 1 shows that 95% lexical coverage can be achieved with knowledge of the 4k to 5k most frequent word families. Since participants' average vocabulary size was 6k (with a minimum size of 4k), these figures suggest that they should not have had any major comprehension difficulties.

Target words

The scripts from the two documentaries were inspected using the corpus software AntConc (Anthony, 2011, Version 3.2.2), from which a set of twenty-six target words (TWs) that were judged by the researchers to be unknown to participants were selected (23 nouns, 2 verbs, and 1 adjective). The judgment process consisted of evaluating and discussing the words with other L2 teachers and deciding whether the TWs were likely to be known by the learners (Peters, et al., 2016). Also, a pilot study conducted with a representative sample indicated that it was unlikely that students were familiar with the words. The target words belonged to the 3,000 most frequent word families and beyond (except for *branch*, which belongs to the 2,000, but it is not commonly used as a verb). The TWs were grouped according to the captioning condition they belonged to (captioned and uncaptioned), as shown in Table 2.

Table 2. *List of target words including word class*

Condition	Target words	Word class
Uncaptioned	Murmuration	Noun
	Smithereens	Noun
	Overlord	Noun
	Quadrotor	Noun
	Dab	Noun
	Eon	Noun
	Gills	Noun
	Abalone	Noun
	Mucin	Noun
	Gripper	Noun
	Slime	Noun
	Swarm	Noun
	Hagfish	Noun
Captioned	Aileron	Noun
	Slat	Noun
	Chum	Noun
	To branch	Verb
	Flop	Verb
	Flaps	Noun
	Stooge	Noun
	Oobleck	Noun
	Denticles	Noun
	Humvee	Noun
	Newtonian	Adj.
	Sharklet	Noun
	Gecko	Noun

Instruments

Vocabulary Pretest

Participants' prior knowledge of the TWs was measured using a vocabulary sight-recognition checklist, which has been used in previous studies and has proven to be a valid format (e.g. Neuman & Koskinen, 1992; Schmitt, Jiang, & Grabe, 2011). However, to avoid participants' over- or underestimating their knowledge, a definition in Spanish was also required. The test contained the 26 TWs and 32 distracters, which were high frequency words with similar spelling to the TWs to make the test more challenging. The scoring procedure was as follows: 1 point was awarded when the definition was fully correct, 0.5 when the definition was correct but incomplete, and 0 when the answer was not related to the word. The reliability coefficient for the checklist was .67 (Cronbach Alpha).

Vocabulary Tests

In order to diminish Pretest sensitization (Swanborn & de Glopper, 1999), i.e. the priming of words from the Pretest to the Posttest, different instruments were used as Posttests: Form-recall/Spelling and Meaning-recognition. The tests were developed on the online platform specialized in test design and administration *Classmarker* (www.classmarker.com). The whole test contained 52 questions divided into two sections: Form-recall/Spelling and Meaning-recognition Test. The platform does not allow resuming the test or going back to previous questions.

The Form-recall/Spelling test contained 26 items with defining context sentences and a blank for the target word to be filled in. To control for possible synonyms to be inserted into the blanks and to give hints about the TWs, one or two initial letters were provided at the beginning of the blank, as suggested by Baltova (1999). This test was used to assess knowledge of two different subcomponents of lexical mastery; thus, two scoring procedures were followed. To measure form recall, accurate spelling was not necessary for a TW to be judged correct. If the word was equal or similar to the target word (e.g. *Avalon* instead of *Abalone*), it was awarded one full credit. If the word was half of the correct word (e.g. *Dents* instead of *Denticles*) or if the word was related to the TW (e.g. *Dermal* instead of *Denticles*, mentioned in the documentary as *Dermal Denticles*), 0.5 points were given. If there was no answer or the answer was incorrect, 0 points were awarded. Participants' knowledge of the spelling of the TWs was gauged using an adapted version of Barcroft's (2002) Lexical Production Scoring Protocol-Written (LPSP-Written): 1 point was given if the word was correctly spelled; .75 if the word had 1 or 2 spelling mistakes (e.g. *gekko* instead of *gecko*); .5 if the spelling was similar to the target word (e.g. *Nweton* instead of *Newtonian*), and 0 points if the word was not similar at all or if it was incorrect. As the same test was analyzed with different criteria, two separate sets of scores were derived from it (Form -recall Test and Spelling Test for the analysis). The Cronbach Alpha coefficient was .76 for form-recall scoring and .77 spelling scoring, indicating a high internal consistency coefficient.

The Meaning-recognition Test was a standard 26 multiple-choice-item test with a stem, three distracters, and an *'I do not know'* option to reduce the effect of guessing (Brown, et al., 2008; Rodgers, 2013). The participants had to choose the meanings that were nearest to the TWs. The options were definitions of the same part of speech as the stem, and the distracters came from

different semantic sets to allow small amounts of knowledge to be demonstrated (Brown, et al., 2008). The words in the definitions belonged to the 2K most frequent word families. Whenever this was not possible, the explanatory words were cognates in the Spanish language. The scoring procedure was as follows: one point was given when the answer was correct. The internal consistency coefficient (KR-20) was .69, which indicates a medium internal reliability.
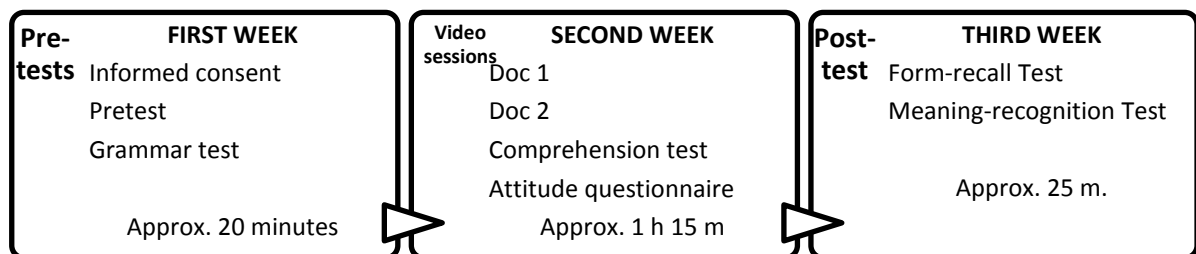
Validity of the instruments was established using different methods, following the recommendations of Alderson, Clapham, and Wall (1995). First, to establish content validity, all the tests were shown to two vocabulary experts to ensure they measured form recall and meaning recognition and that they had only one possible answer. In order to determine its face validity, two native speakers checked that the instruments contained natural and understandable language (Schmitt, 2010), and that the questions could be answered if the words were known. Finally, response validity was established by piloting the materials and testing instruments on a representative sample.

Procedure

Spanish learners of English were contacted through invitations posted in several social networks and through personal connections. Volunteering participants were scheduled to take part in the study. They were informed about the general purpose of the study and signed their informed consent. The viewing sessions took place individually or in small groups (depending on the participants' availability) and the tests were always conducted individually.

The whole experimental phase is represented in Figure 2.

Figure 2. *Data collection procedure*

Pre-tests	FIRST WEEK	Video sessions	SECOND WEEK	Post-test	THIRD WEEK
	Informed consent		Doc 1		Form-recall Test
	Pretest		Doc 2		Meaning-recognition Test
	Grammar test		Comprehension test		
			Attitude questionnaire		Approx. 25 m.
	Approx. 20 minutes		Approx. 1 h 15 m		

Before taking the Pretest, the participants were encouraged not to guess and not to make use of any external source, such as Google or dictionary. A grammar test was given after the Pretest to flush memory. Participants were then scheduled to view the documentaries the following week. The Pretest was given on average a week before the documentary viewing session ($M = 7.24$ days, $SD = 4.95$).

   The participants viewed the documentaries either on their laptops or in the University Library computers, with headsets. The first documentary was presented uncaptioned and the second captioned. They were told that their task was to pay attention to the documentaries because they would have to answer some very specific comprehension questions. They were not allowed to take notes. Additionally, the participants were not informed about the upcoming vocabulary tests to ensure they would not study the unknown words in the text (Hulstijn, 2001). To further ensure this would not happen, the Comprehension Test and the Attitude Questionnaire after the treatment served to distract the learners' attention away from explicit reflection of unknown items (and potential explicit study).

   Approximately ten days after the video session ($M = 10.56$, $SD = 4.86$), the participants were sent a code via e-mail, with which they had to register on the online platform www.classmarker.com to take the Vocabulary Tests. They completed the Vocabulary Tests at home according to their availability. The time allotted was 45 minutes, but the participants took 24' 27" on average. On completion of the tests, they were shown the correct answers and were debriefed about the specific purpose of the study.

Results

The descriptive statistics for the results of the Vocabulary Tests per captioning condition are presented in Table 3.

Table 3. *Descriptive statistics (Mean and Standard Deviation), N = 34*

Condition	Pretest		Form recall		Spelling		Meaning recognition	
Uncaptioned	1.0	(1.37)	1.59	(1.78)	1.38	(1.62)	8.12	(2.04)
Captioned	6.99	(1.09)	2.51	(1.88)	2.07	(1.70)	9.06	(2.17)

In order to answer the research questions (i.e. *How much vocabulary are participants able to learn incidentally from uncaptioned and captioned video viewing, as measured by Form-recall/Spelling and Meaning-recognition Tests?* and *Does captioning lead to higher incidental L2 vocabulary gains than words presented in audio-visual input with no captions, as measured by Form-recall/Spelling and Meaning -recognition Tests?),* a two-way repeated measures ANOVA was run. It helped determine the effect of Tests (four levels: Pretest, form recall, spelling, meaning recognition) and Captioning condition (two levels: captioned, uncaptioned) (independent variables) on the participants' raw scores (dependent variable) and the interaction effect between Tests and Captioning condition. Mauchly's test of sphericity indicated that the

assumption of sphericity had been violated for the two-way interaction, $\chi^2(5) = 48.23$, $p < .0001$, and the Tests, $\chi^2(5) = 94.68$, $p < .0001$. The test for sphericity for the Captioning condition was not computed since it contained only two levels. As a result, the Greenhouse-Geisser coefficient (ε) was interpreted for the main effects and the interaction. For the two-way interactions and simple main effects, a Bonferroni adjustment was applied to the level at which statistical significance was declared. There was a statistically significant two-way interaction between Tests and Captioning condition on the participants' scores, $F(1.96, 64.718) = 5.70$, $p = .006$, $\varepsilon = .65$, but the effect size was small, $\eta^2 = .02$. As the interaction was significant, simple main effects were run.

Results from the simple main effect comparisons showed that there was a statistically significant simple main effect for uncaptioned ($F(1.82, 60.176) = 256.71$, $p < .0001$, $\varepsilon = .61$, $\eta^2 = .75$), as well as for captioned words ($F(1.98, 65.4) = 328.34$, $p < .0001$, $\varepsilon = .66$, $\eta^2 = .77$). In both cases, the effect sizes were strong.

The main effect analysis was followed by pairwise comparisons (post-hoc test) for Tests to determine if there was a statistical difference between the raw scores from the Pretest and Posttests (see Table 4). The analysis showed that in the uncaptioned condition the increase in the scores of only the Meaning- recognition Test were found to be statistically significant. Regarding the captioned condition, the increases in all the tests were statistically significant. This suggests that incidental learning from viewing is possible as far as meaning recognition is concerned; however, form recall and spelling need the presence of on-screen text to take place.

Table 4. *Pairwise comparison of Tests in relation to the Pretest expressed in number of words*

Captioning condition	Comparison with the Pretest	M difference	SD error	p	95% Confidence interval for difference	
					Lower bound	Upper bound
Uncaptioned	Form-recall	.53	.29	.476	-.29	1.35
	Spelling	.32	.28	1.0	-.36	-.07
	Meaning-recognition	7.06***	.33	<.0001	6.14	7.98
Captioned	Form-recall	1.53***	.28	<.0001	.74	2.32
	Spelling	1.08**	.26	.001	.35	1.81
	Meaning-recognition	8.07***	.34	<.0001	7.13	9.02

**p< .001
***p< .001

Table 4 shows that in total there was a rise in the scores from Pre to Posttests of 2.06 words on the Form-recall Test, from which 1.4 were correctly spelled, and 15.13 on the Meaning-recognition Test.

To further explore the role of captions (research question 2), gains per participants were calculated. However, as participants who knew more TWs had fewer chances for improvement than participants who knew fewer TWs, absolute gains (Mean differences, Table 4) were not considered to give a complete picture of the participants' vocabulary learning. As a result, participants' relative gains were used in the analysis. Unlike the absolute gains, relative gains take into account the participants' varying opportunities for an increase in knowledge (Rodgers, 2013). The formula to calculate relative gain percentages was developed by Shefelbine (1990) and subsequently used in different studies on incidental vocabulary acquisition (Horst, et al., 1998; Montero Perez, et. al., 2014; Rodgers, 2013; Webb & Chang, 2015). The formulas are displayed in Figure 3.

Figure 3. *Absolute and relative gains formulas*

$$\text{Absolute gains \%} = \frac{\text{(Posttest scores - Pretest scores)}}{\text{Total number of TW}} \times 100$$

$$\text{Relative gains} = \frac{\text{(Posttest scores - Pretest scores)}}{\text{(Total number of TWs - Pretest scores)}} \times 100$$

Table 5 shows participants' absolute and relative gains expressed in percentages. Descriptive statistics showed that in most cases relative gains were slightly higher than absolute gains.

Table 5. *Absolute and relative gains in % for the two Vocabulary Tests per captioning condition per participant (N = 34; TWs = 26, 13 in the uncaptioned and 13 in the captioned conditions)*

Captioning condition	Tests	Absolute gains		Relative gains	
		M	SD	M	SD
Uncaptioned	Form-recall	4.07	13.12	3.84	14.86
	Spelling	2.43	12.40	1.93	14.25
	Meaning-recognition	54.30	14.66	59.46	15.46
Captioned	Form-recall	11.76	12.63	12.66	13.87
	Spelling	8.31	11.74	8.81	12.77
	Meaning-recognition	62.10	15.15	67.61	17.34

A two-way repeated measures ANOVA was run to determine the effect of Captions (two levels: captioned, uncaptioned) and the different Tests (three levels: form recall, spelling, meaning recognition) (independent variables) on the participants' relative gain scores (dependent variable) and the interaction effect between Tests and Captioning condition. Mauchly's test of sphericity indicated that the assumption of sphericity had been violated for the two-way interaction, $\chi^2(2) = 55.4$, $p < .0001$, and the simple main effect of Tests, $\chi^2(2) = .51$, $p < .0001$. The test for sphericity for the Captioning condition was not computed since it contained only two levels. As a result, the Greenhouse-Geisser coefficient (ε) was interpreted for the interaction and the simple main effects. For the two -way interactions and simple main effects, a Bonferroni adjustment was applied to the level at which statistical significance was declared. There was no statistically significant two-way interaction between Tests and Captioning condition on the participants' scores, $F(1.10, 36.54) = .28$, $p = .625$, $\varepsilon = .55$, $\eta^2 = .00$, which means that the increases were not statistically different across Tests. Therefore, the main effects were interpreted for only the Captioning condition.

The simple main effect of captions showed a statistically significant difference in the participants' relative gains, $F(1,33) = 17.52$, $p < .0001$, $\varepsilon = 1.00$, $\eta^2 = .10$. The effect size was moderate to strong. The Captioned condition ($M = 29.63$, SD error $= 2.09$) was statistically significantly higher than the uncaptioned condition ($M = 21.74$, SD error $= 1.97$), with a mean difference of 7.95 and SD error of 1.90 (95% CI, 4.09 to 11.81) % of words. The analysis showed that scores from the captioned condition were approximately 8% statistically higher than scores from the uncaptioned condition in all the three tests.

Discussion

The aim of this study was to investigate the incidental acquisition of L2 vocabulary through the viewing of two science documentaries and the role that captions played in that process. To that end two two-way repeated measures ANOVA tests were conducted. The first analysis was carried out with the raw scores to determine if there was a significant increase from Pretest to Posttest, whereas the second was performed with the participants' relative gains to look at the effect of captions on incidental vocabulary gains.

In response to the first research question, whether participants were able to learn new words incidentally from documentary viewing, results from this study indicate that it is possible to acquire words incidentally after viewing a science documentary. The significant outcome of the effect of video viewing on incidental vocabulary acquisition is consistent with the earlier findings of Baltova (1999), Huang and Eskey (1999), Hui (2007), Markham (1999), Markham, Peter, and McCarthy (2001), Montero Perez, et al. (2015), Montero Perez, et al. (2014), Neuman and Koskinen (1992), Peters et al. (2016), Sydorenko (2010), Rodgers (2013), Vidal (2003, 2011), Webb and Peters (2018), and Winke, Gass and Sydorenko (2010), who found that video viewing does lead to incidental vocabulary acquisition.

Nevertheless, the increase from Pretest to Posttest does not occur under all conditions. On the Form-recall and Spelling Tests, the increase was only significant when words were captioned. A different picture was seen on the Meaning-recognition Test whose increases were more consistent since they were statistically significant in both captioning conditions. This was expected to a certain extent, due to gains are usually modest in what recall is concerned whereas recognition usually reports higher gains (Schmitt, 2010). This is in line with previous research

on incidental learning from video viewing (Montero Perez, et al., 2014; Peters et al., 2016, and Sydorenko, 2010).

It is noteworthy that, although the mean increase on the Form-recall Test (.53) and Spelling Test (.32) in the uncaptioned condition were not statistically significant, an overall increase of 2.06 words in production (form recall), from which 1.4 were correctly spelled, and nearly 15 words in comprehension (meaning recognition) are figures not to be ignored. A Wilcoxon signed-ranked test revealed that the overall increases from Pre to Posttests on the Form-recall and Spelling Tests were statistically significant when not divided into Captioning condition ($p < .005$). These results indicate that audio-visual viewing enhances the initial process of vocabulary acquisition, namely meaning recognition, but, perhaps, more sophisticated information processing (Hulstijn, 2001) is needed to make learners produce the words learned.

On the bright side, Schmitt (2010) recommends a number of 10 words to be taught explicitly per 1-hour class period, which can be done through elaboration in order to increase, in turn, the learners' vocabulary depth (different dimensions of word knowledge). Regarding the results reported here, recognizing on average 15 words and being able to recall 2 seems to be an accomplishment for words that were not meant to be learned, particularly because incidental learning is slower and more gradual than explicit learning (Schmitt, 2008).

In any case, these figures may be conservative estimates of the amount of vocabulary learning that occurred. There are other factors that have not been considered, which may cause the underestimation of the participants' vocabulary gains. First, this study reports increases in only two dimensions of vocabulary learning of the selected target items. However, participants may have strengthened their vocabulary depth in other dimensions (i.e. form recognition, meaning recall) or words that were not assessed. Rodgers (2013) contends that it is likely that participants increased their vocabulary depth through video viewing, at least for words that they had prior knowledge of, especially words in the first or second frequency bands. The author adds that that kind of knowledge is expected to arise in the form of collocations, aural form, and polysemy (p. 98). Webb and Rodgers (2009b) lend support claiming that the rich context encountered in audio-visual material usually favors the improvement of these components of knowledge. This happens because such contextualized learning situations tend to give learners a strong sense of word meaning and use (Huckin & Coady, 1999).

In response to the second research question, results from this study revealed that a 30-minute documentary viewing session with captioning provides the opportunity to gain 8% more words than when they are uncaptioned. The overarching finding from this study is that video viewing does lead to vocabulary learning, but captioning boosts the gains on meaning recognition and makes gains on form recall and spelling possible. This evidence ties in with previous research that has found that vocabulary learning from input with captions is superior to words learned without captions, as far as recall (Baltova, 1999; Neuman & Koskinen, 1992; Sydorenko, 2010; Winke et al., 2010), spelling (Hui, 2007), and recognition (Huang & Eskey, 1999; Hui, 2007; Neuman & Koskinen, 1992; Markham, 1999; Stewart & Pertusa, 2004; Sydorenko, 2010) are concerned. It is worth noting, however, that although some of those studies have found statistical differences, the effect sizes were not always reported. In the case of this study, the magnitude of the main effect of Captioning condition appears to be strong when gains were computed. In contrast, Montero Perez and Desmet's (2012) in-depth meta-analysis of research on the effect of captions on incidental vocabulary learning showed that although apparently convincing, in most cases, results from the studies analyzed were not statistically significant. Consequently, caution is advised when generalizing the findings.

Conclusions

The results of the present study have provided further evidence of the effectiveness of viewing for incidental vocabulary learning, and they have done so using a more ecologically valid type of authentic material appropriate for adult learners. In other words, the method used can approximate a real context of adult learning. It has also supported previous findings that indicate that captions significantly aid meaning recognition, but they are not substantially influential on form recall and spelling when words are uncaptioned. These results might be particularly important in the educational context, overall where teachers hold the view that captions distract learners from the audio and create text dependence (Hui, 2007). Hui (2007) found that, regarding word meaning, higher proficiency students benefit more from captions whereas low proficiency students benefit more from subtitles than captions. In conclusion, supporting authentic audio-visual material with simultaneous on-screen text is a useful method to learn words incidentally, but, based on Hui's findings, teachers should decide carefully when and how to use them. Also, caution is advised since the evidence in favor of captions is not yet conclusive. The results presented in this study imply that video viewing is a good way to present learners with the words so that they can start recognizing their meaning, but the use of other forms of word learning to complement incidental learning and, thus, facilitate word production is advised.

Shortcomings and limitations

The present study is not exempt from having weaknesses and constraints. The experimental design involved a sample that did not belong to an intact language learning group. Although this is arguably a positive aspect since it eliminates the potential effect of teachers and class-specific prior experiences, the conditions and times in which each participant received the treatment and took the tests varied, which might have resulted in some small variations in the results.

With respect to the instruments, different tests were used as Pre and Posttest. As explained, the reason behind this decision was to eliminate test sensitization; however, that decision may have affected how the Vocabulary Tests correlated to the Pretest (although significant moderate to strong correlation coefficients were found between the instruments, .59 Form-recall and .52 Meaning-recognition). In that sense, better strategies of measurement need to be devised to avoid test sensitization while accurately assessing advancements in vocabulary learning.

Finally, the level of word imagery, i.e., how words are represented with the images, how concrete words are, together with how simultaneously the audio is presented with the imagery and text (if applicable) are factors that might have affected the acquisition of words from viewing. Rodgers (2013) emphasizes that the imagery that accompanies the aural input, among other factors, might contribute to the understanding of audio-visual input, which might also foster the incidental learning of words. Future research should address the role of imagery in this process. These are factors that could not be examined in the present study and that need to be addressed in future research.

References

Alderson, J. C., Clapham, C., and Wall, D. (1995). *Language Test Construction and Evaluation*. Cambridge: Cambridge University Press.

Anthony, L. (2011). AntConc (Version 3.2.2) [Computer Software]. Tokyo, Japan: Waseda University. Available from http://www.antlab.sci.waseda.ac.jp/

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). Ontario Institute for Studies in Education of the University of Toronto: Toronto.

Barcroft, J. (2002). Semantic and structural elaboration in L2 lexical acquisition. *Language Learning*, *52*(2), 323-363.

Brown, R., Waring, R., and Donkaewbua, S. (2008). Incidental vocabulary acquisition from reading, reading-while-listening, and listening to stories. *Reading in a Foreign Language*, *20*(2), 136-163.

Bruton, A., López, M. G., and Mesa, R. E. (2011). Incidental L2 vocabulary learning: An impracticable term? *TESOL Quarterly*, *45*(4), 759–768.

Classmarker. (n.d.). Free Quiz Maker for Professionals [Computer Software]. Retrieved July 14, 2016, from https://www.classmarker.com/

Cobb, T. (n.d.). The Complete Lexical Tutor [Computer Software]. Retrieved from http://www.lextutor.ca/.

Chung, J. M. (1999). The effects of using video texts supported with advance organizers and captions on Chinese college students' listening comprehension: An empirical study. *Foreign Language Annals*, *32*(3), 295–308.

Danan, M. (2004). Captioning and subtitling undervalued language learning strategies. *META*, *49*(1), 67-77.

Day, R. R., and Bamford, J. (1998). *Extensive reading in the second language classroom*. Cambridge, England: Cambridge University Press.

DeCarrico, J. S. (2001). Vocabulary learning and teaching. In M, Celce -Murcia (ed.), *Teaching English as a second or foreign language* (285-299). Boston: Heinle and Heinle.

DocuSERIESEngsub For EducationFuturistic. (2016, May 11). *Material Series | Ep04 Smart Materials Design Technology | Documentary English subtitles*. [Video File]. Retrieved May 20, 2016, from https://www.youtube.com/watch?v=niJ3rjrqqF8

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

Horst, M., Cobb, T., and Meara, P. (1998). Beyond A Clockwork Orange: Acquiring second language vocabulary through reading. *Reading in a Foreign Language*, *11*, 207–223.

Huang, H. C., and Eskey, D. E. (1999). The effects of closed-captioned television on the listening comprehension of intermediate English as a Second Language (ESL) students. *Journal of Educational Technology Systems*, *28*(1), 75-96.

Huckin, T., and Coady, J. (1999). Incidental vocabulary acquisition in a second language: A review. *Studies in Second Language Acquisition, 21,* 181-193.

Hui, W. (2007). The effects of captions on Chinese EFL students' incidental vocabulary acquisition. *Media in Foreign Language Instruction*, *30*(4), 9-16.

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.

Hulstijn, J. H. (2003). Incidental and intentional learning. In C, Doughty, and M. H. Long. (eds.) *The handbook of second language acquisition* (pp. 349-381). Oxford: Blackwell Publishing.

Laufer, B., and Ravenhorst-Kalovski, G. C. (2010). Lexical threshold revisited: Lexical text coverage, learners' vocabulary size and reading comprehension. *Reading in a Foreign Language*, *22*(1), 15.

Markham, P. (1999). Captioned videotapes and second-language listening word recognition. *Foreign Language Annals*, *32*(3), 321–328.

Markham, P., Peter, L. A., and McCarthy, T. J. (2001). The effects of native language vs. target language captions on foreign language students' DVD video comprehension. *Foreign Language Annals*, *34*(5), 439–445.

Mayer, R. E. (2009). *Multimedia learning* (2nd ed.) New York: Cambridge University Press.

Meara, P. (n.d.). V_YesNo. Retrieved January 22, 2017, from http://www.lognostics.co.uk/tools/V_YesNo/V_YesNo.htm

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., and Desmet, P. (2012). The effect of input enhancement in L2 listening on incidental vocabulary learning: A review. In Macaire D., Boulton A. (Eds.), *Languages, Cultures and Virtual Communities* (pp. 153-157)*,* Proceedings of EuroCALL 2010, Procedia Social and Behavioral Sciences, 34.

Montero Perez, M., Peters, E., Clarebout, G., and Desmet, P. (2014). Effects of captioning on video comprehension and incidental vocabulary learning. *Language Learning and Technology 18*(1), 118–141.

Nation, I. S. P. (2006). How large a vocabulary is needed for reading and listening? *The Canadian Modern Language Review, 63*, 59–82.

Nation, I. S. P. (2007). The four strands. *Innovation in Language Learning and Teaching*, *1*(1), 2–13.

Nation, I.S.P. (2001). *Learning Vocabulary in Another Language*. Cambridge: Cambridge University Press.

Neuman, S. B., and 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.

NOVA. (2011, January 19). *Making Stuff*. Retrieved April 07, 2017, from http://www.pbs.org/wgbh/nova/tech/making-stuff.html

NOVA. (2013, October 169). *Making more stuff*. [Video File]. Retrieved May 15, 2016, from http://www.pbs.org/wgbh/nova/tech/making-more-stuff.html

Pellicer-Sánchez, A., and Schmitt, P. (2010). Incidental vocabulary acquisition from an authentic novel: Do things fall apart? *Reading in a Foreign Language*, *22*(1), 31-55.

Peters, E., & Webb, S. (2018). Incidental vocabulary acquisition through viewing L2 television and factors that affect learning. *Studies in Second Language Acquisition*, *40*(3), 551-577.

Peters, E., Heynen, E., and Puimège, E. (2016). Learning vocabulary through audiovisual input: The differential effect of L1 subtitles and captions. *System*, 63, 134-148.

Pigada, M., and Schmitt, N. (2006). Vocabulary acquisition from extensive reading: A case study. *Reading in a Foreign Language*, *18*, 1–28.

Rodgers, M. P. H. (2013). English language learning through viewing television: An investigation of comprehension, incidental vocabulary acquisition, lexical coverage, attitudes, and captions (Unpublished doctoral dissertation). Wellington, NZ: Victoria University.

Rodgers, M. P. H., and Webb, S. (2011). Narrow viewing: The vocabulary in related television programs. *TESOL Quarterly*, *45*(4), 689–717.

Rott, S. (1999). The effect of exposure frequency on intermediate language learners' incidental vocabulary acquisition through reading. *Studies in Second Language Acquisition*, *21*, 589–619.

Schmitt, N. (1997). Vocabulary Learning Strategies. In D. N. Schmitt, & M. McCarthy (Eds.), *Vocabulary Description, Acquisition and Pedagogy* (pp. 199-227).

Schmitt, N. (2008). Instructed second language vocabulary learning. *Language Teaching Research, 12*(3), 329–63.

Schmitt, N. (2010). *Researching vocabulary: A vocabulary research manual*. Basingstoke: Palgrave Macmillan.

Schmitt, N. and Schmitt, D. (2014). A reassessment of frequency and vocabulary size in L2 vocabulary teaching. *Language Teaching*, *47*(4), 484 - 503.

Schmitt, N., Jiang, X., and Grabe, W. (2011). The percentage of words known in a text and reading comprehension. *Modern Language Journal*, *95*(1), 26-43.

Shefelbine, J. L. (1990). Student factors related to variability in learning word meanings from context. *Journal of Literacy Research*, *22*(1), 71–97.

Stewart, M., and Pertusa, I. (2004). Gains to language learners from viewing target language closed-captioned films. *Foreign Language Annals, 37*(1), 438-447.

Swanborn, M., and de Glopper, K. (*1999*). Incidental word learning while reading: A meta-analysis. *Review of Educational Research*, *69*, 261-285.

Sydorenko, T. (2010). Modality of input and vocabulary acquisition. *Language Learning and Technology*, *14*(2), 50–73.

van Zeeland, H., and Schmitt, N. (2013). Incidental vocabulary acquisition through L2 listening: A dimensions approach. *System*, *41*(3), 609–624.

Vandergrift, L. (2007). Recent developments in second and foreign language listening comprehension research. *Language Teaching 40*, 191–210.

Vanderplank, R. (2010). De´ja` vu?A decade of research on language laboratories, television and video in language learning. *Language Teaching, 43*(1), 1-37.

Vidal, K. (2003). Academic listening: A source of vocabulary acquisition? *Applied Linguistics 24*, 56-89.

Vidal, K. (2011). A comparison of the effects of reading and listening on incidental vocabulary acquisition. *Language Learning*, *61*, 219–258.

Waring, R., and Takaki, M. (2003). At what rate do learners learn and retain new vocabulary from reading a graded reader? *Reading in a Foreign Language*, *15*, 130–163.

Webb, S., and Chang, A. C. S. (2015). Second language vocabulary learning through extensive reading with audio support: How do frequency and distribution of occurrence affect learning? *Language Teaching Research*, *19*(6), 667–686.

Webb, S. (2007). The effects of repetition on vocabulary knowledge. *Applied Linguistics*, 28, 46-65.

Webb, S. (2011). Selecting television programs for language learning: Investigating television programs from the same genre. *International Journal of English Studies*, 11 81), 117-136.

Webb, S., and Rodgers, M. P. H. (2009a). Vocabulary demands of television programs. *Language Learning*, *59*(2), 335-366.

Webb, S., and Rodgers, M. P. H. (2009b). The lexical coverage of movies. *Applied Linguistics*, *30*(4), 407–427.

Winke, P., Gass, S., and Sydorenko, T. (2013). Factors influencing the reading of captions by foreign language learners: An eye-tracking study. *The Modern Language Journal, 97*(1), 254-275.

Appendix A

Studies on the influence of captions on incidental vocabulary learning. Mean vocabulary scores presented for each study and condition.

		Condition			
Study	TWs	Captions	Other forms of captions	Subtitles	No captions
Baltova (1999)	30	Obs[a]. M= 21.85 (6.71) Adj. M= 20.51		Obs[a]. M= 14.55 (9.04) Adj. M= 14.95	Obs[a]. M= 11.66 (6.35) Adj. M= 12.60
Huang & Eskey (1999)	16	5.07 (1.67) words			3.47 (1.41) words
Hui (2007)	10	-Word Recognition High proficiency: 7.93 (.98) Low Proficiency: 6.16 (1.21) Spelling High proficiency: 6.38 (1.15) Low Proficiency: 5.03 (1.08) -Word-meaning High proficiency: 5.77 (1.30) Low Proficiency: 4.03 (1.11)		-Word Recognition High proficiency: 7.07 (1.28) Low Proficiency: 5.52 (.77) Spelling High proficiency: 5.47 (1.14) Low Proficiency: 4.03 (.95) -Word-meaning High proficiency: 6.20 (1.06) Low Proficiency: 4.65 (.84)	-Word Recognition High proficiency: 5.17 (1.08) Low Proficiency: 4.37 (1.10) Spelling High proficiency: 4.37 (1.19) Low Proficiency: 3.47 (1.04) -Word-meaning High proficiency: 4.07 (1.05) Low Proficiency: 3.30 (1.12)
Markham (1999)	100	-Oral word recognition 69.3 words (9.08)			-Oral word recognition 59.8 words (11.9)
Montero Perez, Peters, & Desmet (2015)	18	-Form recognition: 7.31 (2.75) -Clip association: 5.31 (3.09) -Meaning recall: 1.08 (1.44) -Meaning recognition: 10.15 (2.12)	Keyword -Form recognition: 9.25 (2.30) -Clip association: 6.58 (2.43) -Meaning recall: 1.83 (1.53) -Meaning recognition: 11.33 (1.92)		

Study	N			
Montero Perez, Peters, Clarebout, & Desmet (2014)	17	-Form recognition: 11.07 (2.54) -Clip association: 8.78 (2.52) -Meaning recognition: .60 (.14) -Meaning recall: .16(.22)	Keyword: -Form recognition: 9.87 (3.17) -Clip association: 7.63 (2.90) -Meaning recognition: .59 (.14) -Meaning recall: .14(.20) Full Caption-Highlighted Keyword: -Form recognition: 10.15 (2.82) -Clip association: 7.85 (2.84) -Meaning recognition: .63 (.17) -Meaning recall: .14 (.19)	-Form recognition: 7.13 (3.26) -Clip association: 5.73 (3.17) -Meaning recognition: .53 (.14) -Meaning recall: .13 (.22)
Neuman & Koskinen (1992)	90	-Form recognition: 73.17% (14.97%) -Form recall: 75.7% (10.4%)		-Word recognition: 64.79% (15.31%) -Word recall: 71.8% (15.27%)
Sydorenko, T. (2010)	28	Video, audio & Captions -Written recognition: 73 % (10%) -Aural recognition: 67% (15%) -Written translation: 36% (11%) -Aural translation: 35% (18%)	Video & Captions -Written recognition: 76% (12%) -Aural recognition: 68% (.06%) -Written translation: 28% (10%) -Aural translation: 24% (12%)	Video & audio -Written recognition: 63 % (10%) -Aural recognition: 69% (.08%) -Written translation: 25% (13%) -Aural translation: .18% (.13%)
Peters, Heynen, and Puimege (2016)		Exp 1 Captions: -Form recognition: 48.2% -Meaning recall: 19.3%	Exp 1 Subtitles -Form recognition: 32.4% -Meaning recall: 19.9%	

	Exp 2	Exp 2
	-Form recognition: 29.1%	-Form recognition: 74.5%
	-Meaning recognition: 17%	-Meaning recognition: 31.8%
	-Form recall: 21.5%	-Form recall: 11.1%

[a]Obs.= Observed
Adj.= Adjusted

_____

Appendix B

Summary of data from research on incidental learning from video viewing

Studies	VKS	Form recognition	Meaning recognition	Meaning recall	Form recall	Spelling	Association
Baltova (1999)	8.3%				27.62%		
Huang and Eskey (1999).			26.7%				
Hui (2007)		60.3%		46.7%	47.9%		
Markham (1999)		64.6%					
Markham, Peter and McCarthy (2001)		61.7%					
Montero Perez, Peters, and Desmet (2015).		47.9%	61.4%	11%			36.8%
Montero Perez, Peters, Clarebout, and Desmet (2014)		58.1%	3.5%	.82%			44.9%
Neuman and Koskinen (1992)		68.9%	60.5%		26.6%		
Rodgers (2013)- Study 2		23.0%	29.6%				
Sydorenko (2010).		69.5%		27.5%			
Peters, Heynen, and Puimège (2016)		Exp 1 42.1%		Exp 1 19.9%			
		Exp 2 27.4%	Exp 2 24.2%		Exp 2 16.2%		