

Studies in Second Language  
Acquisition



CAMBRIDGE  
UNIVERSITY PRESS

**Incidental vocabulary acquisition through viewing L2  
television and factors that affect learning**

Journal:	<i>Studies in Second Language Acquisition</i>
Manuscript ID	SLA-RA-2016-0189.R3
Manuscript Type:	Research Article
Keywords:	vocabulary, viewing, L2 television, frequency of occurrence, cognate

SCHOLARONE™  
Manuscripts

Review

## Abstract

Research has begun to demonstrate that L2 words can be learned incidentally through watching audio-visual materials. Although there are a large number of studies that have investigated incidental vocabulary learning through reading a single text, there are no studies that have explored incidental vocabulary learning through viewing a single full-length TV program. The present study fills this gap. Additionally, three word-related variables (frequency of occurrence, cognateness, word relevance) and one learner-related variable (prior vocabulary knowledge) that might contribute to incidental vocabulary learning were examined. Two experiments were conducted with Dutch-speaking EFL learners to measure the effects of viewing TV on form recognition and meaning recall (Experiment 1) and meaning recognition (Experiment 2). The findings showed that viewing TV resulted in incidental vocabulary learning at the level of meaning recall and meaning recognition. The research also revealed that learning was affected by frequency of occurrence, prior vocabulary knowledge, and cognateness.

Research has shown that English language learners need to know approximately 3,000 word families to understand spoken discourse (van Zeeland & Schmitt, 2013; Webb & Rodgers, 2009a, 2009b) and 8,000-9,000 word families to understand written discourse (Nation, 2006; Schmitt, Jiang, & Grabe, 2011). Given the limited amount of classroom time that can be devoted to learning vocabulary, it is unlikely that all these words will be learned in the classroom (Webb & Nation, 2017). Therefore, it is not surprising that researchers have advocated the potential of extensive reading to boost learners' vocabulary knowledge (Nation, 2015; Schmitt, 2008). Yet, Cobb (2007) found that L2 learners would be unlikely to learn the most frequent 3,000 words through extensive reading. Webb and Rodgers (2009a) suggested that watching L2 television might be another way to increase learners' word knowledge, because within a relatively small amount of television viewing, there are repeated encounters with lower frequency words.

Research has shown that vocabulary can be learned incidentally through reading a single text (Horst, Cobb, & Meara, 1998; Pellicer-Sánchez & Schmitt, 2010; Saragi, Nation, & Meister, 1978). Although there is some evidence that L2 vocabulary might be learned incidentally through watching short, educational video clips (e.g., Montero Perez, Peters, Clarebout, & Desmet, 2014; Neuman and Koskinen, 1992), no study has investigated whether it is possible to incidentally learn new words through watching a single, full-length TV program. If incidental vocabulary learning through L2 viewing is able to fuel L2 vocabulary growth, as Webb (2015) suggests, studies that show the potential for incidental learning through viewing TV programs are needed. The present study aims to fill this gap. Additionally, it would be the first study to investigate how a number of word-related variables (frequency of occurrence, cognateness, word relevance) and one learner-related variable (prior vocabulary knowledge) might contribute to incidental vocabulary learning through viewing one full-length television program.

## Background

### Viewing TV

A survey of media use in the European Union showed that 86% of the people surveyed watch TV almost every day (European Commission, 2014). This figure contrasts sharply with Europeans' reading habits, as 40% of the respondents claimed not to read any books (European Commission, 2002). Canadians and Americans watch TV five times more than they read (Statistics Canada, 1998; United States Department of Labor, 2006). Similar results have been found for foreign language learners. Lindgren and Muñoz (2013) showed that watching subtitled movies was a more important source of out-of-class FL exposure for 10-11 year old foreign language learners than reading books. Similarly, Peters' (in press) survey on EFL learners' out-of-class exposure to English in Flanders indicated that more than 40% of the 79 EFL learners surveyed watch (subtitled) English language TV programs and movies several times a week. However, only one percent of the respondents claimed to be engaged in reading English-language books several times a week.

Exposure to L2 input has been shown to be beneficial for language proficiency as well as vocabulary learning (Lindgren & Muñoz, 2013; Peters, in press). The bulk of research into incidental vocabulary acquisition has focused on exposure to L2 reading texts with the majority of studies showing that vocabulary can be learned incidentally through reading a short text or text excerpt (e.g., Chen & Truscott, 2010; Pellicer-Sánchez, 2016; Rott, 1999; Webb, 2007), a single text or novel (e.g., Day, Omura & Hiramatsu, 1991; Dupuy & Krashen, 1993; Horst, Cobb, & Meara, 1998; Pellicer-Sánchez & Schmitt, 2010; Saragi, Nation, & Meister, 1978; Waring & Takaki, 2003; Zahar, Cobb, & Spada, 2001), or through extensive reading (e.g., Webb & Chang, 2015a, 2015b). Recently, a number of corpus studies

have started to focus on the potential benefits of audio-visual input for incidental vocabulary learning (Rodgers & Webb, 2011; Webb & Rodgers, 2009a, 2009b). One of the greatest benefits of TV is that it provides learners with large amounts of authentic, spoken L2 input (Webb, 2015). Compared to reading (Cobb, 2007), TV programs also have the advantage that low-frequency words occur repeatedly in a relatively small amount of viewing time (Webb & Rodgers, 2009a). Moreover, Rodgers and Webb (2011) found that repeated encounters with low-frequency words were even higher in the case of related TV programs, such as episodes from the same TV program. That is why Webb (2015) proposed that extensive TV viewing inside and outside the classroom should be considered a valuable vocabulary learning method in addition to extensive reading because it “could help to fill the need for greater L2 input” in EFL contexts with limited exposure to L2 input (p.159). Given the wide availability of English language TV programs via DVDs, online media services, and streaming, it should be possible for many EFL learners to have easy access to the L2 spoken input found in TV programs. TV could, thus, be another important learning resource for authentic L2 input. Although the results of the abovementioned corpus studies seem promising, the empirical evidence for the potential benefits of TV viewing for incidental vocabulary acquisition is still limited to the use of short video clips (Koolstra & Beentjes, 1999; Montero Perez et al., 2014; Neuman & Koskinen, 1992) or watching multiple episodes of one TV program (Rodgers, 2013). Unlike in reading, no study has investigated the potential of watching a single, full-length TV program.

There is some evidence that L2 vocabulary might be learned incidentally through watching short video clips. One of the first studies investigating foreign language vocabulary learning through TV viewing was Neuman and Koskinen (1992), in which bilingual children (grade 7 and 8) watched three short clips from a children’s program about science. Their findings revealed that there were large vocabulary learning gains for children who had watched the

clips compared to a control group. Neuman and Koskinen were among the first to emphasize the potential of TV viewing for vocabulary learning. Similarly, d'Ydewalle and his colleagues (d'Ydewalle & Van de Poel, 1999; Pavakanun & d'Ydewalle, 1992) found considerable learning gains for vocabulary after learners had watched a short video clip. In another study involving children (L1 = Dutch; L2 = English), Koolstra and Beentjes (1999) found similar findings that indicated that TV viewing contributed to incidental vocabulary acquisition. Interestingly, they also found that children who watched TV programs in English on a regular basis learned more words than children who watched TV less often.

Most studies investigating TV viewing have used short video clips that were a maximum of 15 minutes long. The participants in these studies were also very often children. An exception is Rodgers' (2013) study that explored incidental vocabulary learning through viewing 10 full-length episodes of one TV program. He found that adult L2 learners learned new words and that the learning gains were comparable to those found in reading studies. However, whether L2 vocabulary is learned incidentally through viewing a single TV program in the same way that L2 words are learned through reading a single text (e.g., Horst, Cobb, & Meara, 1998; Pellicer-Sánchez & Schmitt, 2010; Saragi, Nation, & Meister, 1978; Waring & Takaki, 2003) remains to be examined. Taken together, there appears to be a clear need for research investigating the effects of TV viewing if we are to better understand the role audio-visual input might play in L2 lexical development. In the remainder of this section, we will discuss the role of word-related (frequency of occurrence, cognateness, word relevance) and learner-related factors (prior vocabulary knowledge) that have been shown to play a facilitative role in incidental vocabulary acquisition.

### **Word-related factors and incidental vocabulary acquisition**

### Frequency of occurrence

There is robust evidence that repeated encounters with unknown words in written input contribute to vocabulary learning (Horst, Cobb, & Meara, 1998; Pellicer-Sánchez & Schmitt, 2010; Rott, 1999; Webb, 2007). Most reading studies indicate that considerable learning gains can occur after 8 to 10 encounters. However, different aspects of knowledge might need different numbers of encounters. For example, Pellicer-Sánchez and Schmitt showed that to learn meaning recognition one encounter might suffice, whereas Webb (2007) found that to gain productive knowledge of words a greater number of encounters were needed than to gain receptive knowledge. Recent evidence from an eye-tracking study (Pellicer-Sánchez, 2016) has shown that repeated encounters also result in faster reading times of new words.

Research investigating the effect of frequency in listening studies is far more limited. Vidal (2003) found that frequency of occurrence in spoken text affected word learning positively, but other word-related factors explained more variance. In a follow-up study (Vidal, 2011), this finding was corroborated, but this time it was also revealed that the effect of frequency of occurrence was much smaller in listening compared to reading. Finally, van Zeeland and Schmitt (2013) showed that frequency of occurrence (3, 7, 11, or 15 occurrences) did not affect all aspects of word knowledge (form recognition, grammar, meaning recall) in the same way. They found a weak frequency effect on form recognition and grammar (a significant difference between 3 and 7 occurrences), but only in the immediate posttests. The effect of frequency on meaning recall was unexpected, as there was only an effect of 11 occurrences and no differences between 3, 7 or 15 occurrences.

The effect of frequency of occurrence has also been investigated in two viewing studies. In his longitudinal study, Rodgers (2013) found a small correlation between frequency and word learning when a tough test was used (= multiple choice test with distractors sharing aspects of

the form and meaning with the correct answer). However, he did not find a relationship between frequency and word learning in a sensitive test (= a multiple choice test with distractors not semantically related to correct answer). In a study comparing the effect of L1 subtitles and captions, Peters, Heynen, and Puimège (2016) found that frequency of occurrence contributed positively to the learning gains made through viewing a video clip, but there was reason to think that its effect was related to learners' prior vocabulary knowledge because of a significant interaction between frequency of occurrence and learners' vocabulary size in two tests.

From the studies reviewed here, it seems that compared to reading, the effect of frequency of occurrence through listening and TV viewing is less straightforward. Further research to clarify the role of frequency in audio-visual input, thus, seems warranted.

### **Cognates**

Traditionally, cognates are words that are formally (phonologically or orthographically), semantically, and etymologically related in two languages. In SLA studies, however, the definition that is often adopted is words that are formally and semantically related, e.g. the English word *house* and the German word *Haus*, or the English word *cat* and the Dutch word *kat* (Rogers, Webb, & Tanaka, 2015). Generally, there are more cognates between genetically related languages, such as English and Dutch (Schepens, Dijkstra, Grootjen, & Van Heuven, 2013).

From psycholinguistic research, we know that cognates are easier to learn than non-cognates (de Groot & Keijzer, 2000; Lotto & de Groot, 1998). However, these studies have also been criticized for their lack of ecological validity. This issue was addressed in a recent study by Rogers, Webb, and Nakata (2015) that compared the learning of loanwords and non-cognates



by Japanese EFL learners. Learning gains were measured in two tests, a cloze test and a translation test. The results showed greater learning of the loanwords than the non-cognates in the immediate and delayed translation test. However, learning gains were greater for the non-cognates in a cloze test. So the facilitative effect of cognates was only partially confirmed.

To our knowledge, no study to date has investigated the role of cognates in a TV viewing study. Though it did not specifically target cognate words, a study of Dutch-speaking children by d'Ydewalle and Van de Poel (1999) is relevant because they explored the effect of linguistic similarity on incidental vocabulary learning. Greater learning gains were revealed for Danish than for French, which might be attributed to the larger number of cognates between Danish and Dutch (66%) than between French and Dutch (20%) (Dyren et al., 1992)<sup>1</sup>. In her study on academic listening and reading, Vidal (2011) found greater gains for L2 English words that were similar to L1 Spanish words than words without formal and semantic similarity. Additionally, her findings also showed that the role of cognates was considerably larger in aural input compared to written input. Of the four word-related variables under investigation (frequency of occurrence, type of elaboration, type of word, word form predictability), cognacy was the variable that explained most of the variance. Vidal argued that L2 listeners pay more attention to words that are similar in their L1 as they are probably more salient in speech. She concludes that cognates have “a clearer facilitative effect for listeners” than for readers (p.246). Similarly, Lindgren and Muñoz (2013) suggest that cognates might play a more facilitative role in listening than in reading because cognate linguistic distance (= lexical similarity based on the proportion of cognates) explained more variance in young learners' listening comprehension than in their reading comprehension. Van der Slik (2010) also found that cognate linguistic distance was a greater predictor of speaking proficiency than writing proficiency. Given that cognates might be more salient in

the input, especially in aural input, it is worthwhile to explore their role in TV viewing as well.

### **Word relevance**

Relatively little attention has been paid to *word relevance* in studies of vocabulary acquisition. Moreover, different definitions have been adopted for word relevance. One definition is *task-induced* word relevance, which means that relevant words are words that are essential to answering reading comprehension questions. Task-induced relevance was investigated in two studies (Peters, 2007; Peters, Hulstijn, Sercu, & Lutjeharms, 2009), which both showed that words that were more relevant to completing a task were better retained than words that were not relevant to task completion.

Word relevance can also be defined as words relevant to understanding a text. For instance, Vidal (2003, 2011) looked at the role that different *types of words* (technical, academic, low-frequency words) played in listening and reading. Technical words were defined as words closely related to the topic of the texts and essential to understanding the lectures. These words could thus be considered relevant words. Her findings revealed that the type of word was the second best predictor of word learning through listening, whereas it explained the least variance in reading. Also the best learning gains were found for the technical words that were crucial to understanding the gist of the lecture. This indirectly suggests that words relevant to understanding a listening text are more likely to be noticed and learned. Given the paucity of research into this variable, the present study aims to examine the relationship between relevance and incidental vocabulary learning through TV viewing to shed more light on its role in vocabulary acquisition.

### **Learner-related factors and incidental vocabulary acquisition: Prior vocabulary knowledge**

Learners with a larger vocabulary size tend to understand reading and listening texts better than learners with a smaller vocabulary size (Laufer & Ravenhorst-Kalovski, 2010; Noreillie, Kestemont, Heylen, Desmet & Peters, in press; Stæhr, 2009; Schmitt, Jiang & Grabe, 2011). Similarly, it has been shown that prior vocabulary knowledge also plays a role in incidental vocabulary acquisition. Horst, Cobb, and Meara (1998) revealed that there was a positive correlation between prior knowledge and learning gains, albeit not a strong one. Webb and Chang (2015a) carried out a longitudinal study which focused on vocabulary learning through extensive reading. They also found an effect of prior vocabulary knowledge on incidental vocabulary acquisition, as higher-level participants learned significantly more words than lower-level participants.

The role of prior vocabulary knowledge has been addressed in a limited number of TV viewing studies, but its effect seems to be less consistent than in reading studies. Although Neuman and Koskinen (1992) used an oral proficiency measure rather than a vocabulary test, they found that learners' level of linguistic ability affected learning gains; high-proficiency learners learned significantly more words than mid- and low-proficiency learners. Similarly, in two studies investigating the effects of different types of captions (Montero Perez et al., 2014; Montero Perez, Peters, & Desmet, 2015), a positive relationship between prior knowledge and vocabulary learning by French-as-a-foreign-language learners was revealed. Peters et al. (2016) also found a positive correlation between prior vocabulary knowledge and word learning when EFL learners had watched a subtitled excerpt from a documentary or a cartoon episode. In contrast, Rodgers (2013) did not find that learners with greater vocabulary knowledge learned more words through viewing L2 television than learners with less vocabulary knowledge. A final aim of the present study is to examine the relationship

between prior vocabulary knowledge and vocabulary learning to help clarify the earlier findings.

### **Rationale and research questions**

The present study seeks to investigate whether watching a full-length TV-program affects word learning. It is important to study the potential of audio-visual input because TV provides learners with authentic, spoken input and creates opportunities for incidental vocabulary learning. Previous research has mainly used short video clips to investigate the potential of TV for language learning. Yet, it is important to investigate whether learning occurs in longer TV-programs as well, because such viewing conditions would better reflect authentic viewing behavior. Little is known about the relationship between prior vocabulary knowledge and word learning through TV viewing. Additionally, there is little known about the relationships between frequency of occurrence, cognateness, and relevance and vocabulary learning through TV viewing. Given the importance of these variables to incidental vocabulary acquisition through reading and listening, more research into their role in TV viewing is necessary. This study is the first to examine the potential of watching one full-length TV program for word learning. Moreover, it aims to shed light on a number of variables that might influence the learning process.

The following research questions were addressed:

1. Is there an effect of viewing a full-length episode of a L2 TV program on word learning?

2. What is the relationship between word learning through viewing a full-length episode of a L2 TV program and the following variables: frequency of occurrence, relevance, cognateness, and prior vocabulary knowledge?

To answer those questions, two experiments using different test instruments were conducted. Both studies adopted a pretest-posttest-delayed posttest between-participants design. English-as-a-Foreign-Language learners (EFL) with Dutch as their L1 were randomly assigned to either an experimental or control group. The experimental group was exposed to the audio-visual input, whereas the control group was not. Experiment 1 measured the effect of TV viewing on form recognition and meaning recall, whereas Experiment 2 measured meaning recognition. Nation and Webb (2011) emphasized the importance of measuring different degrees of vocabulary knowledge. This is why we tested three aspects of word knowledge (form recognition, meaning recall, and meaning recognition). However, Nation and Webb also warned of potential test effects. The reason we set up two experiments was first to avoid a possible test effect, while still testing three word knowledge aspects. Secondly, we also wanted to reduce the potential for test fatigue among participants because taking a meaning recall as well as meaning recognition test, each containing 64 items, might take too much time and be demotivating for participants. The documentary, the questionnaire, the tests, the target items, and the procedure were piloted with a group of participants resembling the participants in the two experiments. On the basis of the pilot results, the number of target items was reduced (see also below). No other changes were made. We will first present the methodology and results of Experiment 1 (form recognition, meaning recall), before moving on to Experiment 2 (meaning recognition).

## Experiment 1: The effect of TV viewing on form recognition and meaning recall

### Method

#### Participants

Sixty-three Flemish EFL business students (L1 = Dutch) in their first or second year at university took part in Experiment 1. Participants had an intermediate proficiency level (B1-B2 according to the *Common European Framework of Reference [CEFR]*). It should be noted that the participants in this study should have been used to watching TV in English, because research (Peters, in press) has shown that in addition to subtitled TV programs, approximately 40% of the Flemish EFL university students surveyed also reported to watch TV programs and movies without subtitles on a weekly basis.

The participants were randomly assigned to either an experimental group, who watched the TV program, or a control group, who only took the tests. There were 36 participants in the experimental group and 27 in the control group. We only included data of participants who attended at least the pretest session and the posttest session, which explains why the number of participants in the two groups is different.

#### Audio-visual input

An authentic full-length TV program that was relevant to the participants' course objectives (Business English course) was selected as the material. Unlike earlier studies that used short video clips that were a maximum of 15 minutes long, the TV program we selected was a full-length, one-hour BBC documentary on the economist J.M. Keynes. The documentary was piloted with a group of participants resembling the participants in this experiment. Findings

from a questionnaire showed that learners found the documentary interesting and relevant. Additionally, the questionnaire results of the pilot also indicated that the level of difficulty was appropriate. This was further corroborated in an open question asking learners what they had learned in terms of content. We were, thus, confident that the documentary would be appropriate for our target population. An analysis of the lexical profile of the documentary using RANGE (Nation & Heatley, 2002) and Nation's (2012) BNC/COCA word lists revealed that 90.3% of the vocabulary was from the most frequent 2,000 word families, and 2.57% were from the 3,000 word level.

### Target items

One of the reasons reported for low learning gains in incidental vocabulary acquisition through reading research is the small number of target items (Pellicer-Sánchez & Schmitt, 2010). To avoid this, 64 target items were selected from the documentary (see Table 1). In the pilot, 111 potential target words were tested. Items that were known by 80% or more of the participants were no longer included in the final target item selection. We also removed polysemous words (e.g. *to foster*, *vicious*) because these items are difficult to score in meaning recall tests. Table 1 lists the items and their frequency of occurrence (FoO), cognate status, and perceived relevance (see also below).

Table 1

*Target items with frequency of occurrence in input, cognate status, and relevance*

Item	FoO	Cognate	Relevance	Item	FoO	Cognate	Relevance
bubble	5	Yes (3/3)	4.3	foresee	2	Yes (2/3)	5.25

slump	5	No (0/3)	6.15	grapple	2	No (0/3)	3.88
treaty	6	No (0/3)	5.08	thrive	2	No (0/3)	4.75
treasury	6	No (0/3)	4.63	alliance	1	Yes (3/3)	5.25
wage	5	No (0/3)	4.63	bursar	1	No (0/3)	3.25
beggar	5	No (0/3)	6.25	conflagration	1	No (0/3)	6.25
impose	5	No (0/3)	4.3	impoverishment	1	No (0/3)	6.75
tame	5	Yes (2/3)	5.5	trait	1	No (0/3)	3.25
boom	4	Yes (2/3)	6.25	lever	1	No (0/3)	4.25
austerity	3	No (0/3)	6.75	riot	1	No (0/3)	3.5
bonds	3	No (0/3)	5.67	sanctuary	1	No (0/3)	2.25
cure	3	No (1/3)	4.75	scrounger	1	No (0/3)	5
currency	3	No (0/3)	6.5	stature	1	No (0/3)	4.25
debtor	3	No (0/3)	6.42	swathe	1	No (0/3)	1.75
deficit	3	Yes (2/3)	6.33	wheat	1	No (0/3)	2.75
herd	3	No (0/3)	4.58	greedy	1	No (0/3)	4.25
legacy	3	No (0/3)	4.67	adamant	1	No (0/3)	3
policy	3	No (0/3)	4.67	onerous	1	No (0/3)	6
tire	3	No (0/3)	1.33	adverse	1	No (0/3)	4.5
boost	3	Yes (2/3)	5	ferocious	1	No (0/3)	2.5
pave	3	No (0/3)	3.75	delineate	1	No (0/3)	4.5
persuade	3	No (0/3)	3.83	demolish	1	No (0/3)	2.5
step in	3	No (0/3)	4.42	ease	1	No (0/3)	4.25
solar	3	No (0/3)	1.58	endorse	1	No (0/3)	4.25
broadcast	2	No (0/3)	1.25	hail	1	No (0/3)	2.25
bust	2	No (1/3)	6.25	inherit	1	No (0/3)	3.25
commodity	2	No (0/3)	4.5	mend	1	No (0/3)	6
dole	2	No (0/3)	4.63	relish	1	No (0/3)	3



monetary	2	Yes (3/3)	6.63	slash	1	No (0/3)	4.75
profound	2	No (0/3)	3.25	stack	1	No (0/3)	5.5
queue	2	No (0/3)	2.38	steer	1	Yes (2/3)	6
successor	2	No (0/3)	3.75	totter	1	No (0/3)	5

---

### Variables

The relationship between word learning through TV viewing and the following four variables: frequency of occurrence, cognateness, relevance, and learners' prior vocabulary knowledge, were also examined.

**Frequency of occurrence:** Frequency of occurrence of the target items was one of the word-related variables taken into account in this study. Frequency of the target words ranged from 1 (e.g. *impoverishment*) to 6 occurrences (e.g. *a treaty*). Because the documentary was created for L1 viewers in the UK and beyond, the frequency of occurrence of the items had ecological validity.

**Cognateness:** To investigate the effect of cognate versus non-cognate items, three raters who were fluent in both English and Dutch were asked to indicate whether a target word was a cognate word with Dutch or not. Items that were considered cognates by at least two raters were labeled cognate items (e.g. *an alliance – een alliantie*), whereas items that were considered cognates by no raters or only one rater were labeled non-cognate items.

**Relevance:** The second variable was a target word's relevance to understanding the content of the documentary. Our definition of word relevance is similar to that used in Vidal (2003, 2011) in that the words were rated for their relevance to comprehension, thus words rated higher were perceived to be more useful for understanding the documentary while those rated lower were perceived to be less useful. Word relevance was operationalized by having three raters assess the relevance of the target items to understanding that passage on a seven-point-scale, with one "being not relevant to understanding the passage" and seven "being very relevant to understanding the passage". We used the average score of the three raters in our analyses (see also Table 1)<sup>2</sup>. The raters were all graduate students in an applied linguistics program in Canada. All of the raters had experience teaching ESL or EFL.

**Learners' prior vocabulary knowledge:** Learners' vocabulary knowledge was measured by means of a frequency-based multiple choice meaning recognition test (Peters, Velghe, & Van Rompaey, 2015). The test, which consists of 120 words, provides an estimate of learners' vocabulary knowledge at different frequency levels. Items were taken from four frequency bands from the COCA frequency lists (Davies, 2008): 30 items from the first 2,000 words, 30 items from the third 1,000 words, 30 items from the fourth 1,000, and 30 items from the fifth 1,000 words. The items are presented in isolation and are accompanied by five options in English: 1 solution, 3 distracters, and 1 "I don't know"-option. The test had good internal consistency (Cronbach's alpha = .95; N = 63) and showed an implicational scale whereby scores on sections related to lower-frequency words were lower than scores for high-frequency words (see also Results section).

Example of test item

*Amazing*

- Very good
- Not very important
- Including many details
- Behaving in an angry way
- I don't know the answer

**Test instruments**

Learners' knowledge of the target items was measured in a paper-and-pencil test consisting of two parts. One part focused on form recognition, the second part on meaning recall. The same test was used as pretest, posttest, and delayed posttest. The items were presented in both their written and aural forms

In the form recognition portion of the test, participants had to tick off whether they could recognize the word. The words, which were audio-recorded, were read twice by a native speaker of English. The order of the items was the same in the pretest and posttests to ensure that the aural forms of the words would be the same. After answering an item in the form recognition test, learners were asked whether they could give the meanings of the items in the meaning recall portion of the test. Learners could provide a translation, synonym or definition. This means that the participants heard and saw the word *to stack*, for instance, had to tick off *yes* or *no* to the question *Have you ever heard the word before?* and subsequently had to provide its meaning if they could. This format minimized the test duration, in comparison to administering two separate tests.

Example of form recognition and meaning recall test

Item	Have you ever heard the word before? Do you recognize the word?	Give the meaning (translation, English synonym, definition, ...)
to stack	<input type="checkbox"/> yes <input type="checkbox"/> no	
adverse	<input type="checkbox"/> yes <input type="checkbox"/> no	
to persuade	<input type="checkbox"/> yes <input type="checkbox"/> no	

Learners were also asked to complete a questionnaire (see Appendix 1). The questionnaire consisted of ten five-point-scale questions and two open questions. The closed questions tapped into learners' perception of the input. One open question focused on what learners had learned in terms of content. This question was asked to verify whether learners had understood the gist of the documentary. The second open question focused on what participants had learned in terms of language, grammar, vocabulary. This allowed us to find other learning gains that were not measured in the tests we used.

### Procedure

The data was collected in three sessions. One week before the experimental treatment, all participants took the pretest and the prior vocabulary knowledge test. In the second session, participants were randomly assigned to either the experimental or the control group and completed an informed consent form. The experimental group watched the documentary on Keynes, which was shown without any subtitles or captions (= L2 subtitles). Next, these participants filled in the questionnaire about the documentary and what they thought they had learned before they took the unannounced posttests. The experimental treatment (video, questionnaire, posttests) took approximately 1 hour and 45 minutes. One week later, students

assigned to the experimental group were tested on their knowledge of the target items without any warning. The control group was not exposed to the audio-visual input. They only took the prior vocabulary knowledge test, the pretest and immediate posttest. All participants were debriefed about the aims of the study.

### Scoring and analyses

All data was scored dichotomously with 0 for an incorrect response and 1 for a correct response. The meaning recall tests were scored by two raters. The interrater reliability was .99 for both the pretest and the posttest. As data was normally distributed, an analysis of covariance (ANCOVA) with learners' prior vocabulary knowledge as covariate was computed to answer the first research question. A repeated measures logistic regression or Generalized Estimating Equation (GEE) in SPSS was conducted to determine which word-related and learner-related variables explained the learning gains in the posttests. Only the data of the experimental group was used in this analysis. A GEE has the advantage that both item-related and participant-related parameters can be combined in one model, which is not the case in an ANOVA or multiple regression analysis. The analysis is based on the number of cases and not on total test scores or total learning gains per participant. This means that the combination "participant, item, response" defines for each observation a particular score (correct/incorrect) on a particular item for a particular participant. For each parameter, the odds ratio ( $= \exp(B)$  or exponential parameter estimate) is calculated that predicts the odds of a correct response. We entered the following parameters into the model: learners' prior vocabulary knowledge, frequency of occurrence, cognateness, and word relevance. Non-significant parameters were always removed from the model before the model was refit.

## Results Experiment 1 (Form recognition, meaning recall)

### Prior vocabulary knowledge test

The descriptive statistics for the vocabulary knowledge test are shown in Table 2. Internal reliability of the vocabulary knowledge test was high (Cronbach's alpha = .95). The results of the test indicated that both groups had mean scores that would indicate mastery of the most frequent 2,000 words (28/30). The participants' mean scores for the most frequent 3,000 words approached mastery (25/30). A *t*-test revealed that the two groups did not differ significantly from each other in terms of prior vocabulary knowledge,  $t = 1.15$ ,  $df = 61$ ,  $p = .25$ ,  $d = .29$ .

Table 2

*Mean scores and standard deviations (in brackets) per group and test section*

Group	Total score (Max = 120)	2K (Max = 30)	3K (Max = 30)	4K (Max = 30)	5K (Max = 30)
Experimental ( $N = 36$ )	91.72 (15.87)	27.89 (2.25)	25.17 (4.15)	21.03 (4.90)	17.64 (5.71)
Control ( $N = 27$ )	96.33 (15.44)	28.04 (2.16)	25.33 (3.83)	23.11 (4.31)	19.85 (6.09)

**Research question 1: Is there an effect of viewing a full-length episode of a L2 TV program on word learning?**

### *Form recognition*

The Cronbach's alpha values were .86 in the pretest and .91 in the posttest. A word was considered learned when it was not known in the pretest, but known in the posttest (= absolute gains). Words known in both the pretest and posttest were considered known words, but not learned words. We used learners' relative gains to determine whether the two groups differed significantly, because relative gains take learners' scores of target items on the pretest into account.

$$\text{Relative gains} = (\text{absolute gains} / \text{number of target items} - \text{number of known words}) \times 100$$

As can be seen in Table 3, both groups performed better on the immediate posttest than on the pretest.

Table 3

#### *Descriptive statistics for form recognition test*

Group	Pretest (Max = 64) Mean ( <i>SD</i> )	Immediate Posttest (Max = 64) Mean ( <i>SD</i> )	Absolute gains Mean ( <i>SD</i> )	Relative gains (%) Mean ( <i>SD</i> )
Experimental ( <i>N</i> = 36)	38.11 (8.67)	43.06 (10.07)	5.58 (5.63)	23.30 (25.12)
Control ( <i>N</i> = 27)	38.96 (7.36)	44.12 (8.43)	5.54 (6.29)	21.24 (22.85)

The ANCOVA with prior vocabulary knowledge as covariate showed that there was no difference between the two groups,  $F(1, 59) = 0.05$ ,  $df = 1$ ,  $p = .83$ ,  $\eta_p^2 = .001$ , and that

learners' prior vocabulary knowledge did not affect learners' relative learning gains,  $F(1, 59) = .76$ ;  $df = 1$ ,  $p = .39$ ,  $\eta_p^2 = .013$ . This was likely due to a learning effect from taking the pretest. The test format asked participants if they had heard the target words before. This creates some ambiguity because when taking the posttest, the participants in both groups had heard the target words before, on the pretest, as some participants informally told one of the researchers. As a result, the data in the form recognition posttest cannot be considered valid and will not be analyzed any further.

### *Meaning recall*

As can be seen in Table 4 and Figure 1, both groups made progress from the pretest to the immediate posttest.

Table 4

*Descriptive statistics for meaning recall pretest, posttest, absolute and relative gains*

Group	Pretest (Max = 64) Mean (SD)	Immediate posttest (Max = 64) Mean (SD)	Absolute gains Mean (SD)	Relative gains (%) Mean (SD)
Experimental (N = 36)	15.67 (8.07)	19.53 (8.47)	3.97 (3.31)	8.31 (6.72)
Control (N = 27)	16.96 (8.05)	18.11 (8.25)	1.56 (1.45)	3.35 (3.03)



Learners in the experimental group learned four words on average (increase of 8.31%), whereas learners in the control group learned 1.5 words (increase of 3.35%). An ANCOVA with prior vocabulary knowledge as covariate revealed that the relative gains were significantly larger for the experimental group than for the control group,  $F(1, 60) = 15.99$ ,  $df = 1$ ,  $p < .0001$ ,  $\eta_p^2 = .21$ . Our treatment “watching TV” explained 21% of the variance (Norouzian & Plonsky, 2017; Plonsky, 2015). The analysis also showed that learners’ prior vocabulary knowledge affected learning significantly,  $F(1, 60) = 5.51$ ,  $df = 1$ ,  $p = .02$ ,  $\eta_p^2 = .08$ . Prior vocabulary knowledge explained 8% of the variance (Norouzian & Plonsky, 2017; Plonsky, 2015).

INSERT FIGURE 1 ABOUT HERE

#### *Delayed meaning recall test*

Although the delayed posttest was administered to the participants, unfortunately the results of this test cannot be attributed solely to the learning condition. The reason for this is that in the delayed meaning recall test, we found a relatively large number of newly learned items. There were 142 correct responses, of which 59 items were not correct in the immediate test. Given this relatively large number of newly learned items, the learning gains found on the delayed posttest might have been the result of deliberate learning between the immediate and delayed posttest. Although from a learning perspective, it is positive that the tests drew learners’ attention to new words, it is problematic from a research perspective. Given that these results may not have been due to only watching the TV program, these findings will not be discussed further.

**Research question 2: What is the relationship between word learning through viewing a full-length episode of a L2 TV program and the following variables: frequency of occurrence, relevance, cognateness, and learners' prior vocabulary knowledge?**

To answer the second research question, a repeated measures logistic regression analysis or GEE in SPSS was carried out with the data from the immediate meaning recall test for the experimental group. The analysis was computed for 1740 observations, i.e. there were 1740 cases, in which the item was not known in the pretest and could potentially be learned (see also Table 5).

Table 5

*Number/percentage of incorrect/correct responses in the immediate meaning recall posttest*

	Incorrect responses	Correct responses	Responses
Experimental	1552	188	1740
group	89.2%	10.8%	100%

The analysis revealed that three parameters contributed significantly to the model: prior vocabulary knowledge, frequency of occurrence, and cognateness (see Table 6). For each of these three variables, there was a positive relationship with word learning. The odds ratio values showed the following.

1. For each additional correct response in the prior vocabulary knowledge test, the odds of a correct response on the posttest were 3% higher. With 10 more words known, the odds of a correct response on the posttest were 32% higher ( $\exp^{10 \cdot B} = 1.32$ ).

2. For each additional occurrence with a target item in the television program, the odds of a correct response were 25% higher ( $\exp(B) = 1.25$ ). For every five additional occurrences, the odds were three times higher ( $\exp^{5*B} = 3.05$ ).
3. When a target item was a cognate, the odds of a correct response were eight times higher.

Word relevance did not contribute significantly to the model.

Table 6

*GEE: Immediate meaning recall test*

Parameter	Wald Chi-square	df	Sig	B	Exp(B)	CI	
						Lower	Upper
Intercept	75.48	1	<.0001	-5.52	.004	.001	.014
Vocabulary	17.17	1	<.0001	.028	1.028	1.02	1.04
FoO	24.51	1	<.0001	.223	1.250	1.15	1.37
Cognate <sup>3</sup>	157.35	1	<.0001	2.193	8.962	6.36	12.63

*Note: Vocabulary = prior vocabulary knowledge; FoO = frequency of occurrence; cognate = cognateness; Exp(B) = odds ratio; CI = Confidence interval for Exp(B)*

### Experiment 2: The effect of TV viewing on meaning recognition

To avoid a test learning effect from pretest to posttest, the effect of viewing TV on meaning recognition was investigated in a second experiment with different participants. It is not unlikely that participants taking a meaning recall test might learn the meaning of some test

items when taking a meaning recognition test immediately after the meaning recall test because of the multiple choice options, in which one option is the correct meaning. That is why we decided to test meaning recognition in a second experiment with different participants.

The same audio-visual input (Keynes documentary), 64 target items, and procedure as described in Experiment 1 were used. However, a different test (meaning recognition instead of form recognition and meaning recall) was administered to the participants.

## **Method**

### **Participants**

Sixty-two Dutch-speaking EFL learners participated in Experiment 2. These participants were also recruited from the first and second year of a Business Administration program. They were considered to be at a CEFR B1/B2 proficiency level. The participants were randomly assigned to the two groups; 37 participants were in the experimental group and 25 were in the control group. The different number of participants in each group can be explained by absences of some students, because only data of participants that attended the pretest as well as the posttest session was included.

### **Test instruments**

Learners' knowledge of the target items was measured in a meaning recognition test. This test used a multiple-choice format. The items were presented in isolation and were accompanied by four options in English: one solution, two distracters, and one "I don't know"-option to minimize guessing. The distracters were definitions of other words that were

used in the television program. As in Experiment 1, the target items were presented in their spoken and written form.

Example of an item in meaning recognition test

A debtor

- a person or company that agrees to do work or provide goods for another company
- a person, group, or organization that owes money
- someone who gets money or something by asking other people for it rather than by paying for it themselves
- I don't know

The same questionnaire as in Experiment 1 was used.

## Results

The scoring procedures and analyses were identical to the ones described in Experiment 1.

### Prior Vocabulary knowledge test

The analyses of the vocabulary test (Cronbach's alpha = .96) showed that the two groups were familiar with the most frequent 2,000 words and most of them were also familiar with the most frequent 3,000 words (see Table 7). A *t*-test revealed that the two groups did not differ significantly from each other in terms of prior vocabulary knowledge,  $t = -1.03$ ,  $df = 60$ ,  $p = .31$ ,  $d = .26$ .

Table 7

*Mean scores and standard deviations (in brackets) per group and test section*

Group	Total score (Max = 120)	2K (Max = 30)	3K (Max = 30)	4K (Max = 30)	5K (Max = 30)
Experimental ( <i>N</i> = 37)	94.11 (18.19)	28.16 (2.10)	25.65 (4.08)	22.08 (5.16)	18.22 (7.93)
Control ( <i>N</i> = 25)	89.40 (16.93)	27.96 (2.49)	24.32 (4.89)	19.64 (4.70)	17.48 (5.90)

**Research question 1: Is there an effect of TV viewing on word learning?**

As can be seen in Table 8, learners in the experimental group were able to recognize almost 14% more words in the posttest than in the pretest (see also Figure 2). An ANCOVA with prior vocabulary knowledge as covariate indicated that the relative learning gains of the experimental group were significantly larger than for the control group,  $F(1, 61) = 5.42$ ,  $df = 1$ ,  $p = .02$ ;  $\eta_p^2 = .084$ . Viewing TV accounted for 8% of the variance in the meaning recognition test (Norouzian & Plonsky, 2017; Plonsky, 2015), which is less than in the meaning recall test in Experiment 1 (21%). Furthermore, prior vocabulary knowledge also had a significant effect on learning,  $F(1, 61) = 5.51$ ,  $df = 1$ ,  $p < .0001$ ,  $\eta_p^2 = .19$ . It explained 19% of the variance in the meaning recognition posttest, which is larger than in the meaning recall test in Experiment 1 (8%). Reliability of the meaning recognition tests ranged from .91 to .92.

Table 8

*Descriptive statistics for meaning recognition test*

Group	Pretest (Max = 64) Mean (SD)	Immediate posttest (Max = 64) Mean (SD)	Absolute gains Mean (SD)	Relative gains (%) Mean (SD)
Experimental (N = 36)	30.95 (10.26)	34.68 (11.36)	3.97 (3.44)	13.95 (13.46)
Control (N = 27)	29.64 (8.88)	29.92 (10.46)	1.68 (2.25)	5.95 (9.33)

Insert Figure 2 about here

***Delayed tests***

Similar to the results of Experiment 1, more than half of the correct responses in the delayed meaning recognition test were newly learned items (= items not known in the immediate posttest). Compared to the meaning recall test, the number of newly learned items was greater because some participants might have learned some items while taking the immediate posttest, as they were exposed to the correct meaning. It is also likely that some participants looked up words at home. Finally, guessing, being inherent in multiple choice tests (Gyllstad, Vilkaite & Schmitt, 2015), might have played a role as well. As the knowledge shown in the delayed meaning recognition test cannot solely be attributed to the viewing treatment, these data will not be discussed any further.

**Research question 2: What is the relationship between word learning and the following variables: learners' prior vocabulary knowledge, frequency of occurrence, relevance, and cognateness?**

The analysis, which was run for 1221 observations (= items unknown in the pretest) (see also Table 9), showed that the following three parameters had a significant and positive correlation with the immediate posttest scores (see Table 10): prior vocabulary knowledge, frequency of occurrence and cognateness, which is in line with the findings from Experiment 1. The factor relevance did not contribute significantly to the model.

Table 9

*Number/percentage of incorrect/correct responses in meaning recognition posttest*

	Incorrect responses	Correct responses	Total
Experimental	922	299	1221
group	75.5%	24.5%	100%

When the score in the test of prior vocabulary knowledge increased by 1, the odds of a correct response in the posttest were 3% higher. With an increase of 10 words, the odds of a correct response were 32% higher ( $\exp^{10 \cdot B} = 1.32$ ). Secondly, the odds in favor of a correct response were 20% higher when frequency of occurrence increased. Five occurrences more resulted in odds that were twice as high ( $\exp^{5 \cdot B} = 2.46$ ). Finally, the odds of learning an item were 2.5 times higher with a cognate word, which is lower than in Experiment 1.

Table 10



*GEE: Immediate meaning recognition test*

Parameter	Wald chi-square	df	Sig	B	Exp(B)	CI	
						Lower	Upper
Intercept	27.22	1	<.0001	-4.14	.016	.003	.08
Vocabulary	13.84	1	<.0001	.028	1.029	1.01	1.04
FoO	14.36	1	<.0001	.180	1.197	1.09	1.31
Cognate	19.89	1	<.0001	.924	2.520	1.68	3.78

*Note: Vocabulary = prior vocabulary knowledge; FoO = frequency of occurrence; cognate = cognateness; Exp(B) = odds ratio; CI = Confidence interval for Exp(B)*

### Summary

Viewing TV had a significant effect on meaning recall (Experiment 1) as well as meaning recognition (Experiment 2). Additionally, the following three parameters had a positive relationship with word learning: learners' prior vocabulary knowledge, frequency of occurrence, and cognateness. Frequency of occurrence had a slightly larger impact on meaning recall than on meaning recognition. However, cognateness seemed to affect meaning recall much more than meaning recognition. The effect of prior vocabulary knowledge was similar in the two tests. There was no correlation between relevance and word learning.

### Discussion

This study expands on earlier research in two ways. It is the first study to investigate vocabulary learning through viewing a single full-length TV program. It is also the first

viewing study to focus on the relationship between frequency of occurrence, cognateness, word relevance, learners' prior vocabulary knowledge and word learning. Taking into account the small incidental vocabulary learning gains that typically occur through reading (Pellicer-Sánchez & Schmitt, 2010), a large number of target items found in the TV program were measured to provide a more accurate measurement of incidental learning. Additionally, strength of vocabulary knowledge was taken into account by measuring form recognition, meaning recall, and meaning recognition.

### **Is there an effect of viewing TV on word learning?**

The first research question can be answered positively. The findings in Experiment 1 and Experiment 2 show that watching a single, full-length TV program can result in substantial learning gains at the level of meaning recall and meaning recognition. However, no learning gains were found in the form recognition test (Experiment 1). This may perhaps have been due to a test effect where knowledge of the forms of target words was gained through taking the pretest. The findings, thus, add to the growing body of evidence that suggests that incidental vocabulary acquisition through viewing TV does occur.

Both participants in Experiment 1 and in Experiment 2 learned approximately four words after watching a one-hour TV program, which corresponds to relative learning gains of 8% in the meaning recall test and almost 14% in the meaning recognition test. This is an encouraging finding. Rodgers (2013) revealed a learning gain of six words through viewing 10 episodes of a TV program. However, unlike the participants in this study, the majority of his participants had not mastered the most frequent 2,000 words, which might have played a role. Additionally, Rodgers argued that the learning gains in his study might have been an underestimation of gains because only words with 5 occurrences in the 3K-14K range were

selected. This was probably the case in the present study as well, as not all words with more than one occurrence were tested. In the questionnaire, the participants in the experimental group were asked what they had learned after watching the TV program. All participants referred to the content, which indirectly shows that they at least understood the gist of the TV program, and some also referred to the use and pronunciation of words, pointing to an increase in learners' depth of knowledge. Participants also listed words that were not tested, such as *accelerate, growth, loan, rescue plan, recession, depression (the great depression), hyperinflation, animal spirits, prosperous, enigma, multiplier, budget cuts, Keynesianism, production plant, the New Deal, and shares*. Clearly, in addition to the words tested, participants learned other words and aspects of knowledge of other words as well.

Previous research (European Commission, 2012; Koolstra & Beentjes, 1999; Lindgren & Muñoz, 2013; Peters, in press) has shown that EFL learners watch a lot of TV and that they do this more often than reading. The present study indicates that watching TV can be “an effective method of learning vocabulary” (Webb & Rodgers, 2009a, p. 356). The findings support Webb's (2015) suggestion that there may be value in viewing of L2 TV inside the classroom where comprehension can be supported by teachers, and that once students are able to enjoy watching L2 TV, autonomous viewing outside the classroom should be encouraged. Extensive viewing of L2 television may thus be a useful complement to extensive reading in developing lexical knowledge.

**What is the relationship between word learning and the following variables: frequency of occurrence, relevance, cognateness and learners' prior vocabulary knowledge?**

#### **Frequency of occurrence**

A positive relationship between frequency of occurrence and vocabulary learning was found in this study. To the best of our knowledge, frequency of occurrence has only been investigated in two TV viewing studies: in the context of viewing multiple episodes of a television program (Rodgers, 2013) and in combination with captions and L1 subtitles (Peters et al., 2016), making direct comparisons difficult. The study by Peters et al. (2016) showed a positive relationship between frequency of occurrence and vocabulary learning, whereas Rodgers (2013) only found a medium-sized correlation ( $r = .30$ ) in one of two tests used, viz. the tough vocabulary test (= multiple choice test in which the distractors shared aspects of form and meaning with the correct answer). Altogether, these three studies tentatively suggest that repeated encounters in L2 TV with new words might facilitate vocabulary learning. It appears that when a word occurs more frequently, it is more likely to be noticed and to become familiar.

The current study showed that repetition had a slightly bigger impact on meaning recall than on meaning recognition. Earlier reading studies (Pellicer-Sánchez & Schmitt, 2010; Webb, 2007) have also demonstrated that different aspects of vocabulary knowledge might need a different number of encounters. Unlike most research into reading, frequency of occurrence was not a factor that was manipulated or controlled for in our study. Slightly more than half of the target items occurred more than once. The highest number of occurrences was six (*a treaty, Treasury*). Six items occurred 5 times; 16 items occurred 3-4 times; and 11 items occurred twice. In the Keynes documentary, there were a number of word types that occurred numerous times (e.g. *economy* (92), *economies* (32), *economic* (60), *world* (24), *global* (80)), but that were not selected as target items as they were either 1K or 2K words or words shown to be familiar in the pilot.

At the same time, our findings also indicate that frequency of occurrence might not be the most important predictor, thus lending evidence to Vidal's (2003, 2011) findings. Although

there was a positive correlation, it might be less important in aural input compared to written input (Vidal, 2011) because L2 learners might encounter more problems in properly decoding and segmenting the speech compared to written input. So they might need more encounters when listening for frequency to have an effect on word learning (Vidal, 2011). However, it is not clear whether repetition has the same impact on listening as on TV viewing. As mentioned, the effect of frequency of occurrence was slightly larger in the meaning recall test than in the meaning recognition test. Yet, frequency of occurrence in spoken input had little effect on meaning recall in van Zeeland and Schmitt's (2013) study, despite more repetitions with target items. In that study, frequency of occurrence only affected immediate meaning recall after 11 encounters. It is not easy to directly compare the two studies given the many methodological differences (non-words vs. real words; four short listening passages vs. one-hour documentary; different L1's vs. one L1), but it could be that frequency of occurrence plays a different role in TV viewing than in listening because of the visual support.

### **Cognateness**

Of the four variables investigated, cognateness was the parameter with the largest effect. The odds of learning a cognate compared to a non-cognate were eight times higher in the case of meaning recall and 2.5 times higher in the case of meaning recognition. Our findings seem to lend evidence to claims made earlier (Lindgren & Muñoz, 2013; Vidal, 2011) that cognateness might play a more central role in aural than in written input. Vidal argues that L2 learners pay more attention to cognates because they are salient in aural input. Given that learners cannot go back to previous words in spoken input as they can with written input, they probably rely more on words similar to L1 words.

The positive findings in this study might also be explained by the fact that Dutch, the participants' L1, and English, the target language, are related languages sharing a high number of cognates. The effect of linguistic similarity on vocabulary learning through TV viewing was already demonstrated in D'Ydewalle and Van de Poel's (1999) study, in which Dutch-speaking children learned more Danish than French words due to greater similarities between Danish and Dutch vocabulary. More recently, Van der Slik (2010) -in a large-scale study of the role of learners' L1 on speaking and writing proficiency- showed that cognate linguistic distance (= lexical similarity based on the proportion of cognates) explained 60 to 40 % of the variance and was more important in speaking than in writing, mirroring the findings of studies on receptive skills (Lindgren & Muñoz, 2013; Vidal, 2011). Although more research investigating the impact of learners' L1 is necessary, it is clear that L1-L2 overlap can play a facilitative role in learning new words.

### Relevance

One unanticipated finding was the lack of a relationship between relevance and vocabulary learning. Unlike Vidal (2011), who found greater learning gains for technical words that were important to understanding the gist of a lecture, we did not find any effect of relevance. It should be pointed out that the technical words in Vidal's (2011) study were closely related to the topic of the input, whereas in the present study relevance was determined by three raters' assessment of the relevance of each of the target items to understanding the documentary on a 7-point scale. The documentary we used also contained words that were closely related to the topic, but words such as *economic*, *economy*, *global* were not included in the final selection of the target items because they did not meet the selection criteria (> 2K words and not known by 80% or more of the participants in the pilot study).

The input modality could also offer an explanation why our findings differ from Vidal (2011). L2 listeners can only rely on speech, whereas L2 viewers can also rely on visual imagery. Additionally, it might be that learners' prior vocabulary knowledge is more important than relevance, but this was not explored in Vidal's study. Obviously, this interpretation remains speculative at this point, and further research is needed to support this hypothesis.

### **Prior vocabulary knowledge**

The findings showed that there was a positive relationship between learners' prior vocabulary knowledge and vocabulary learning through viewing TV. Irrespective of the aspect of knowledge tested (meaning recall or meaning recognition), the effect of prior knowledge was the same. The odds of learning new words while viewing TV were larger for learners who knew more words. These results corroborate previous viewing studies (Montero Perez et al., 2014; 2015; Peters et al., 2016) that took learners' prior vocabulary knowledge into account.

A lexical analysis of the documentary had shown that the most frequent 2,000 words provided approximately 90% text coverage, meaning that participants who knew these words would be able to understand 90% of the running words in the documentary. The most frequent 2000 words from the BNC/COCA lists included a small number of high frequency proper nouns such as *America*, *Americans*, *Europe*, *European*, *German*, *Germans*, and *Germany*, but did not include other proper nouns included in the documentary such as *John*, *Maynard*, *Keynes*, and *Keynesianism*. The results of the present study suggest that learners who knew more words were more likely to gain knowledge of unknown words than those who knew fewer words. This in turn suggests that learners with greater lexical coverage of the spoken text encountered in television, were more successful in incidentally learning

words encountered in the documentary. Given that the vocabulary test (lemmas) and the Range program (word families) used different units of counting, it is difficult to determine the exact amount of lexical coverage of the participants. However, on average, the participants' lexical coverage was likely close to 95%, because they scored high on the 2K and 3K frequency bands of the vocabulary test. Webb and Rodgers (2009a) suggested that learners are likely to need a vocabulary size of 2,000 to 4,000 word families (plus proper nouns and marginal words) in order to understand L2 TV. They argue that if learners regularly view L2 TV, there is great potential for vocabulary learning. The findings of this study provide some support for that hypothesis.

### **Limitations and suggestions for future research**

It should be noted that the participants in this study were used to being exposed to authentic English input, because TV programs and movies in Flanders are subtitled and not dubbed. EFL learners in Flanders watch TV programs and movies with and without subtitles a few times a week (Peters, in press), although the former occurs more frequently. Koolstra and Beentjes (1999) also found a positive relationship between vocabulary learning and the frequency with which children watched subtitled TV programs. Given the omnipresence of English in Flanders via TV, movies, and the Internet, it should be borne in mind that the findings of this study might not be generalizable to all EFL learning contexts.

It should also be noted that the learning gains found in this study might underestimate the extent of learning. This is because target items that were in the first two-thousand most frequent words of English or that were known by 80% of the participants in the pilot were not selected. However, it is possible that these words were learned to some degree. Although partial vocabulary learning was taken into account, it would also have been useful to measure



other aspects of word knowledge. Research suggests that measuring multiple aspects of vocabulary knowledge is likely to provide a more accurate evaluation of learning (e.g., Webb, 2005, 2009a, 2009b). Additionally, follow-up interviews (Pellicer-Sánchez & Schmitt, 2010) could provide a more complete and fine-grained picture of the learning gains.

One aim of this study was to investigate the role of a number of factors shown to be important in incidental vocabulary acquisition. However, it would be useful to examine the role of visual imagery in a future study. Imagery can help viewers in learning the meaning of new words. However, the support that visual imagery actually provides may differ between genres (Rodgers, in press).

Although delayed posttests were administered, the study is limited by the fact that it was not possible to provide clear evidence of the long-term effects of watching a full-length TV program. There were learning gains between the immediate and delayed tests. However, the gains could not solely be attributed to the learning treatment. Webb, Newton, and Chang (2013) reported a similar problem in their study. This seems to cast, at least, some doubt about the validity of delayed posttests in pretest, immediate posttest, delayed posttest designs. It has been previously argued that if the focus is on initial form-meaning mapping, immediate tests may suffice, because repeated encounters as well as repeated retrievals are necessary for a word to be entrenched in the mental lexicon (Hulstijn, 2003; Nation & Webb, 2011). Additionally, Nation and Webb (2011) have warned that when there are low scores, a few participants looking up words can seriously affect delayed posttest scores. It would be useful for follow up studies to investigate the effects of viewing a single L2 television program on vocabulary learning at longer intervals than used in the present study.

### **Conclusion**

This paper reports on two experiments that were carried out to investigate the potential of TV viewing to contribute to vocabulary learning. Both experiments show that incidental vocabulary acquisition through watching a full-length TV program can occur. Participants learned approximately four words at the level of meaning recall as well as at the level of meaning recognition. Both experiments support earlier findings indicating that L2 words are learned incidentally through watching video, and contribute additional evidence that suggests 1) that vocabulary learning is likely to occur through viewing more ecologically valid TV programs 2) that learners' prior vocabulary knowledge affects their learning gains, 3) that there is a positive relationship between frequency of occurrence and learning, and 4) that there is a facilitative effect for cognates. The results suggest that in addition to reading and listening, TV viewing can also be a fruitful method for increasing a learner's vocabulary knowledge. Because people like watching TV, it offers great potential for language learning in the long run. What is now needed, however, is more research into vocabulary learning through viewing. Not only should the study be replicated in other EFL contexts, but there are many other factors whose role in viewing should be further explored.

#### Endnotes

1. Figures for Flemish-Dutch were selected.
2. In our analyses, we first converted the relevance parameter into a dichotomous variable because our assumption was that it would make it easier to interpret and report the data. However, the analyses were also run with relevance as a continuous variable (= using the 7-point-scale). No changes in results were found. The factor relevance did not contribute to the regression model, not as a dichotomous variable and not as a continuous variable. This holds true for the meaning recall (experiment 1) as well as the meaning recognition

test (experiment 2). Given that the outcomes for both types of analyses were the same, we decided to only report the continuous variable. As one reviewer pointed out, changing a continuous variable into a dichotomous one results in loss of data.

3. Reference value is cognates (value = 1). Non-cognates had the value “0”.

## References

- Chen, C. & Truscott, J. (2010). The effects of repetition and L1 lexicalization on incidental vocabulary acquisition. *Applied Linguistics*, 31(5), 693-713.  
<https://doi.org/10.1093/applin/amq031>
- Cobb, T. (2007). Computing the vocabulary demands of L2 reading. *Language Learning & Technology*, 11(3), 38–63. Retrieved from  
<http://llt.msu.edu/vol11num3/cobb/default.html>
- de Groot, A. & Keijzer, R. (2000). What is hard to learn is easy to forget: the roles of word concreteness, cognate status, and word frequency in foreign-language vocabulary learning and forgetting. *Language Learning*, 50(1), 1–56. doi:10.1111/0023-8333.00110/abstract
- Day, R. R., Omura, C., & Hiramatsu, M. (1991). Incidental EFL vocabulary learning and reading. *Reading in a Foreign Language*, 7, 541–551.
- Dupuy, B., & Krashen, S. D. (1993). Incidental vocabulary acquisition in French as a foreign language. *Applied Language Learning*, 4, 55–63.
- D’Ydewalle, G., & Van de Poel, M. (1999). Incidental foreign-language acquisition by children watching subtitled television programs. *Journal of Psycholinguistic Research*, 28(3), 227–245. doi:10.1023/A:1023202130625

- Dyen, I., Kruskal, J.B., & Black, P. (1992). An Indo-European classification: A lexicostatistical experiment. *Transactions of the American Philosophical Society*, 82, 1-132.
- European Commission (2002). Europeans' participation in cultural activities. Retrieved from [http://ec.europa.eu/public\\_opinion/archives/ebs/ebs\\_158\\_en.pdf](http://ec.europa.eu/public_opinion/archives/ebs/ebs_158_en.pdf)
- European Commission (2012). *First European survey on language competences*. Retrieved from: [http://ec.europa.eu/languages/library/studies/executive-summary-eslc\\_en.pdf](http://ec.europa.eu/languages/library/studies/executive-summary-eslc_en.pdf).
- European Commission (2014). Media use in the European Union. Retrieved from [http://ec.europa.eu/public\\_opinion/archives/eb/eb82/eb82\\_media\\_en.pdf](http://ec.europa.eu/public_opinion/archives/eb/eb82/eb82_media_en.pdf)
- Gyllstad, H., Vilkaite, L., & Schmitt, N. (2015). Assessing vocabulary size through multiple-choice formats: Issues with guessing and sampling rates. *ITL International Journal for Applied Linguistics*, 2, 276–303. doi:10.1075/itl.166.2.04gyl
- Horst, M., Cobb, T., & Meara, P. (1998). Beyond a clockwork orange: acquiring second language vocabulary through reading. *Reading in a Foreign Language*, 11(2), 207–223
- Hulstijn, J. H. (2003). Incidental and intentional learning. In C. J. Doughty & M. H. Long (Eds.), *Handbook of second language acquisition* (pp. 349–381). Malden, MA: Wiley–Blackwell
- Koolstra, 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. doi:10.1007/BF02299476
- Laufer, B., & Ravenhorst-Kalovski, G. (2010). Lexical threshold revisited: Lexical text coverage, learners' vocabulary size and reading comprehension. *Reading in a Foreign Language*, 22(1), 15–30

- Lindgren, E., & Muñoz, C. (2013). The influence of exposure, parents, and linguistic distance on young European learners' foreign language comprehension. *International Journal of Multilingualism*, 10(1), 105–129. doi:10.1080/14790718.2012.679275
- Lotto, L., & de Groot, A. (1998). Effects of learning method and word type on acquiring vocabulary in an unfamiliar language. *Language Learning*, 48(1), 31–69. doi:10.1111/1467-9922.00032
- 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. doi:10.1111/modl.12215
- Montero Perez, M., Peters, E., Clarebout, G., & Desmet, P. (2014). Effects of captioning on video comprehension and incidental vocabulary learning. *Language Learning & Technology*, 18(1), 118–141. Retrieved from <http://llt.msu.edu/issues/february2014/monteroperezetal.pdf>
- Nation, I.S.P. (2006). How large a vocabulary is needed for reading and listening? *Canadian Modern Language Review*, 500, 59–82. doi: <http://dx.doi.org/10.3138/cmlr.63.1.59>
- Nation, I. S. P. (2012). *The BNC/COCA word family lists*. Retrieved from <http://www.victoria.ac.nz/lals/about/staff/paul-nation>
- Nation, P. (2015). Principles guiding vocabulary learning through extensive reading. *Reading in a Foreign Language*, 27(1), 136–145
- Nation, I.S.P., & Heatley, A. (2002). Range: A program for the analysis of vocabulary in texts [software]. Downloadable from <http://www.victoria.ac.nz/lals/staff/paul-nation/nation.aspx>
- Nation, I. S. P., & Webb, S. (2011). *Researching and analyzing vocabulary*. Boston: Heinle

- 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), 94–106
- Noreillie, A.-S., Kestemont, B., Heylen, K., Desmet, P., & Peters, E. (in press). Vocabulary knowledge and listening comprehension at an intermediate level in English and French as foreign languages: an approximate replication study of Stæhr (2009). *ITL International Journal of Applied Linguistics*, 169(1).
- Norouzian, R., & Plonsky, L. (forthcoming). Eta- and partial eta-squared in L2 research: A cautionary review and guide to more appropriate usage. *Second Language Research*. <https://doi.org/10.1177/0267658316684904>
- Pavakanun, U., & d'Ydewalle, G. (1992). Watching foreign television programs and language learning. In F. L. Engel, D. G. Bouwhuis, T. Bosser, & G. d'Ydewalle (Eds. ), *Cognitive modelling and interactive environments in language learning* (pp. 193—198). Berlin: Springer.
- Pellicer-Sánchez, A. (2016). Incidental L2 vocabulary acquisition from and while reading. *Studies in Second Language Acquisition*, 38(1), 97–130. doi:10.1017/S0272263115000224
- Pellicer-Sánchez, A., & Schmitt, N. (2010). Incidental vocabulary acquisition from an authentic novel: do Things Fall Apart? *Reading in a Foreign Language*, 22(1), 31–55
- Peters, E. (2007). Manipulating L2 learners' online dictionary use and its effect on L2 word retention. *Language Learning & Technology*, 11(2), 36–58. Retrieved from <http://llt.msu.edu/vol11num2/pdf/peters.pdf>
- Peters, E. (in press). The effect of out-of-class exposure to English language media on learners' vocabulary knowledge. *ITL International Journal of Applied Linguistics*, 169(1).

- Peters, E., Hulstijn, J. H., Sercu, L., & Lutjeharms, M. (2009). Learning L2 German vocabulary through reading: The effect of three enhancement techniques compared. *Language Learning*, 59(1), 113–151. doi:10.1111/j.1467-9922.2009.00502.x/full
- Peters, E., Heynen, E., & Puimège, E. (2016). Learning vocabulary through audiovisual input: The differential effect of L1 subtitles and captions. *System*, 63, 134–148. <http://doi.org/10.1016/j.system.2016.10.002>
- Peters E., Velghe T., Van Rompaey T. (2015, May). *A post-entry English and French vocabulary size for Flemish learners*. Paper presented at EALTA. Copenhagen, Denmark.
- Plonsky, L. (2015). Statistical power, *p* values, descriptive statistics, and effect sizes. A “back-to-basics” approach to advancing quantitative methods in L2 research. In L. Plonsky (Ed.), *Advancing quantitative methods in second language research* (pp. 23-45). New York/London: Routledge.
- Rodgers, M. P. H. (2013). *English language learning through viewing television: An investigation of comprehension, incidental vocabulary acquisition, lexical coverage, attitudes, and captions*. Unpublished PhD thesis, Victoria University, Wellington.
- Rodgers, M.P.H. (in press). The images in television programs and the potential for learning unknown words: the relationship between on-screen imagery and vocabulary. *ITL International Journal of Applied Linguistics*, 169(1).
- Rodgers, M. P. H., & Webb, S. (2011). Narrow Viewing: The Vocabulary in Related Television Programs. *TESOL Quarterly*, 45(4), 689–717. <http://doi.org/10.5054/tq.2011.268062>
- Rogers, J., Webb, S., & Nakata, T. (2015). Do the cognacy characteristics of loanwords make them more easily learned than noncognates? *Language Teaching Research*, 19(1), 9–27. doi:10.1177/1362168814541752

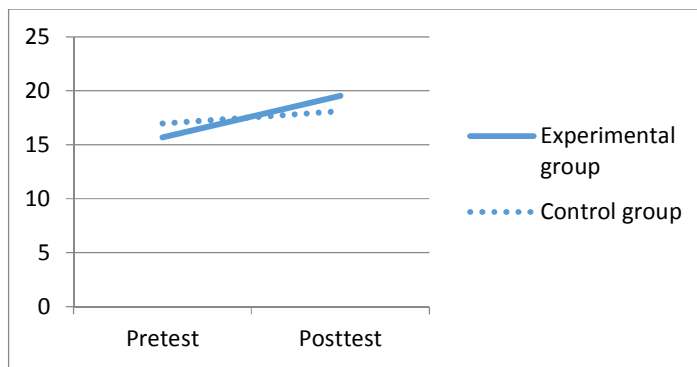
- Rott, S. (1999). The effect of exposure frequency on intermediate language learners' incidental vocabulary acquisition and retention through reading. *Studies in Second Language Acquisition*, 21, 589–619.
- Saragi, T., Nation, I.S.P. & Meister, G.F. (1978). Vocabulary learning and reading. *System*, 6, 72-78.
- Schepens, J., Dijkstra, T., Grootjen, F., & van Heuven, W. J. B. (2013). Cross-language distributions of high frequency and phonetically similar cognates. *PLoS ONE*, 8(5). doi:10.1371/journal.pone.0063006
- Schmitt, N. (2008). Review article: Instructed second language vocabulary learning. *Language Teaching Research* (Vol. 12). doi:10.1177/1362168808089921
- Schmitt, N., Jiang, X., & Grabe, W. (2011). The percentage of words known in a text and reading comprehension. *The Modern Language Journal*, 95(1), 26–43. doi:10.1111/j.1540-4781.2011.01146.x
- Stæhr, L. S. (2009). Vocabulary knowledge and advanced listening comprehension in English as a foreign language. *Studies in Second Language Acquisition*, 31 (4), 577-607. doi:10.1017/S0272263109990039
- Statistics Canada (1998). Average time spent on activities, by sex. Ottawa, Ontario: Author. Retrieved October 21, 2007 from <http://www40.statcan.ca/101/cst01/famil36a.htm>
- United States Department of Labor (2006). American time use survey summary. Washington, D.C.: Author. Retrieved October 21, 2007 from <http://www.bls.gov/news.release/atus.nr0.htm>
- Van der Slik, F. W. P. (2010). Acquisition of Dutch as a second language. *Studies in Second Language Acquisition*, 32(03), 401–432. doi:10.1017/S0272263110000021



- van Zeeland, H., & Schmitt, N. (2013). Incidental vocabulary acquisition through L2 listening: A dimensions approach. *System*, *41*(3), 609–624. doi:10.1016/j.system.2013.07.012
- Vidal, K. (2003). Academic listening: A source of vocabulary acquisition? *Applied Linguistics*, *24*(1), 56–89. doi:10.1093/applin/24.1.56
- Vidal, K. (2011). A Comparison of the effects of reading and listening on incidental vocabulary acquisition. *Language Learning*, *61*(1), 219–258. doi:10.1111/j.1467-9922.2010.00593.x
- Waring, R., & Takaki, M. (2003). At what rate do learners learn and retain new vocabulary from reading a graded reader? *Reading in a Foreign Language*, *15*(2), 130-163.
- Webb, S. (2007). The effects of repetition on vocabulary knowledge. *Applied Linguistics*, *28*(1), 46–65. doi:10.1093/applin/aml048
- Webb, S. (2015). Extensive viewing: language learning through watching television. In D. Nunan & J.C. Richards (Eds.) *Language learning beyond the classroom* (pp. 159-168). New York: Routledge.
- Webb, S., & Chang, A. C. S. (2015a). How does prior word knowledge affect vocabulary learning progress in an extensive reading program? *Studies in Second Language Acquisition*, *37*(4), 651–675. doi:10.1017/S0272263114000606
- Webb, S. & Chang, A, C-S. (2015b). Second language vocabulary learning through extensive reading: How does frequency and distribution of occurrence affect learning? *Language Teaching Research*, *19*(6), 667-686.
- Webb, S. & Nation, I.S.P. (2017). *How vocabulary is learned*. Oxford: Oxford University Press.
- Webb, S., Newton, J., & Chang, A. C. S. (2013). Incidental learning of collocation. *Language Learning*, *63*(1), 91–120. <http://doi.org/10.1111/j.1467-9922.2012.00729.x>

- Webb, S., & Rodgers, M. P. H. (2009a). Vocabulary demands of television programs. *Language Learning*, 59(2), 335–366. <http://doi.org/10.1111/j.1467-9922.2009.00509.x>
- Webb, S., & Rodgers, M. P. H. (2009b). The lexical coverage of movies. *Applied Linguistics*, 30(3), 407–427. <http://doi.org/10.1093/applin/amp010>
- Zahar, R., Cobb, T., & Spada, N. (2001). Acquiring vocabulary through reading: Effects of frequency and contextual richness. *The Canadian Modern Language Review* 57, 541-572. <https://doi.org/10.3138/cmlr.57.4.541>

For Peer Review



*Figure 1.* Interaction between Time and Treatment in Meaning recall test

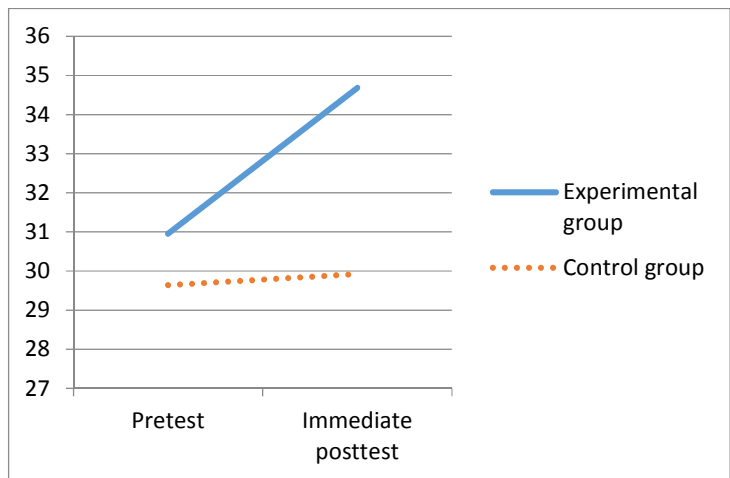


Figure 2. Interaction between Time and Treatment in Meaning recognition test

**Questionnaire**

**1 = strongly DISAGREE – 5 = strongly AGREE**

- |   |   |   |   |   |   |
|---|---|---|---|---|---|
| 1. The topic of the video was interesting.  | 1 | 2 | 3 | 4 | 5 |
| 2. The topic of the video is relevant to this English course.   | 1 | 2 | 3 | 4 | 5 |
| 3. The length of the video was appropriate/OK.  | 1 | 2 | 3 | 4 | 5 |
| 4. The video was easy to understand.  | 1 | 2 | 3 | 4 | 5 |
| 5. I mainly focused on the content of the video.  | 1 | 2 | 3 | 4 | 5 |
| 6. Watching a video is a good way to improve your English.  | 1 | 2 | 3 | 4 | 5 |
| 7. Watching a video is a good way to improve your listening skills in English.  | 1 | 2 | 3 | 4 | 5 |
| 8. Watching a video is a good way to improve your English vocabulary.   | 1 | 2 | 3 | 4 | 5 |
| 9. Watching a video is a good way to improve your grammar skills in English.  | 1 | 2 | 3 | 4 | 5 |
| 10. I paid attention to the words that were used in the video.  | 1 | 2 | 3 | 4 | 5 |
| 11. What have you learned in terms of content?  |   |   |   |   |   |
| 12. What have you learned in terms of language/vocabulary (new words)/grammar? Write as many items as you can remember. |   |   |   |   |   |