Routledge
Taylor & Francis Group

Check for updates

Learning L2 vocabulary from audiovisual input: an exploratory study into incidental learning of single words and formulaic sequences

Eva Puimège and Elke Peters

Research Group Language, Education and Society, KU Leuven, Belgium

ABSTRACT

Television is considered an important source of comprehensible input for second language learners of English and there is some evidence that L2 words can be learned incidentally by watching television. Few studies have looked at the role of TV viewing for learning formulaic sequences, despite the ubiquity of formulaic sequences in spoken English, and the importance of formulaic language in the development of second language proficiency. This study aims to find out whether single words and formulaic sequences can be learned incidentally by watching English language television, and whether learners' prior vocabulary knowledge and item-related factors affect the learning process. Data were collected from 20 English as a foreign language (EFL) learners (L1 = Dutch) in their first year at university. A pre-test post-test within-subject design was adopted. Learning gains were measured at three levels of sensitivity: a form recall test, a meaning recall test and a form recognition test. The results indicate that single words and formulaic sequences can be learned incidentally from TV viewing and that the learning is mediated by item- as well as learner-related factors.

Introduction

Previous research has shown that learners should ideally know the most frequent 8000–9000 word families for reading (Nation 2006) and 3000 for TV viewing (Webb and Rodgers 2009). Because only so much time can be spent teaching vocabulary in the classroom, researchers have encouraged incidental vocabulary learning through exposure to second language (L2) input as a means to expand learners' vocabulary knowledge (e.g. Pellicer-Sánchez 2017; Peters 2018; Peters and Webb 2018; Peters et al. 2019). Incidental vocabulary learning has primarily been researched in the context of reading (e.g. Horst, Cobb and Meara 1998; Pellicer-Sánchez and Schmitt 2010; Saragi, Nation and Meister 1978). However, the increasing popularity of television and movies has incited scholars to investigate the potential of audiovisual input as a source of L2 vocabulary (Webb and Rodgers 2009). Although some recent studies (Montero Perez et al. 2014; Peters and Webb 2018) have shown that words can be learned incidentally through watching TV, the research into vocabulary learning from audiovisual input is still scarce.

Most studies into incidental vocabulary acquisition have focused on single words (Nation 2001). However, little is known about the incidental learning of formulaic sequences (FS). A few studies have shown that FS can be learned incidentally through reading a single text (Choi 2017; Macis 2018; Pellicer-Sánchez 2017) and through reading-while-listening (Webb, Newton and Chang

www.

2013). Yet it remains unclear what the potential of TV viewing is for learning FS. This is surprising, given that a corpus study showed that the distribution of FS on television is similar to that in everyday speech (Lin 2014). The study suggests that television might be a good source of L2 FS, particularly for those FS occurring in spoken discourse. However, there is as yet no conclusive evidence that learners can pick up FS from audiovisual input.

This article reports on an exploratory study that investigated the incidental learning of single words and FS through TV viewing. The first aim was to explore whether FS, like single words, can be picked up incidentally. Secondly, the study aimed to explore some factors that might mediate the learning process of single words and FS.

Background

Incidental learning from audiovisual input

Because learners need to acquire large vocabularies for L2 reading, listening and viewing, it has been argued that L2 vocabulary acquisition should to some extent occur outside the classroom (Peters 2018; Peters et al. 2019). Although most studies on incidental vocabulary learning have been carried out in the context of reading (Nation 2001), there are a number of reasons to consider the role of television as a source of L2 vocabulary. First, viewing is a very popular activity among the majority of L2 learners. A 2017 survey of the European Union (2018) shows that television is the preferred medium among EU citizens, with 84% watching television almost every day. Watching English-language TV is also a popular activity among EFL learners (Peters 2018; Peters et al. 2019). Peters and colleagues found that EFL learners spend considerably more time watching television than reading in English. Moreover, due to its entertainment value, L2 television may be more effective in lowering learners' anxiety than written input, given that learners find TV viewing more accessible than written texts (Larsen-Freeman 1983; Neuman and Koskinen 1992). A third reason why watching television might be beneficial for vocabulary learning is that vocabulary demands for TV viewing (3000 word families) are lower than for reading (Webb and Rodgers 2009).

One of the earliest studies on incidental vocabulary learning from audiovisual input (Neuman and Koskinen 1992) examined whether children could pick up English vocabulary by watching short educational videos with and without captions, and through reading-while-listening. Higher learning gains were found for both viewing groups than for the reading-while-listening group. A recent study by Peters and Webb (2018) examined incidental learning from watching an entire one-hour episode of a TV programme. They found that incidental vocabulary learning through watching TV is possible at the level of meaning recall and meaning recognition. Further, it was shown that word-related factors (cognateness, frequency of occurrence) as well as learners' prior vocabulary knowledge were positively correlated with learning.

Learning formulaic sequences

The term 'formulaic sequence' encompasses a broad range of word combinations and expressions, such as collocations, idioms, phrasal verbs, proverbs and lexical bundles. Two major approaches have been used to define and categorise formulaic sequences: the phraseological approach and the distributional or frequency-based approach (Granger and Paquot 2008). The phraseological approach identifies formulaic sequences based on linguistic criteria, such as syntactic fixedness and semantic opacity. The frequency-based approach, on the other hand, is based on corpus-derived measures such as the frequency and association strength of word combinations. Because previous research has found that both distributional (González Fernández and Schmitt 2015) and phraseological (Peters 2016) properties may affect the learning burden of FS, we combine both approaches in our methodology and loosely define FS as conventionalised, recurring word combinations.

FS constitute up to 60% of spoken and written discourse (Erman and Warren 2000), fulfil many pragmatic functions and make up virtually all conventionalised speech acts (Schmitt 2010). Yet, despite the ubiquity and prominence of FS in language, it has been shown that L2 learners struggle with the appropriate use of FS. Corpus studies have revealed that even advanced learners tend to overuse or misuse FS, and produce many errors, often caused by L1 interference (Laufer and Waldman 2011). Furthermore, L2 learners' knowledge of FS has been shown to lag behind their knowledge of single words (Bahns and Eldaw 1993; Granger 1998; Nesselhauf 2003, 2005).

One reason why L2 learners struggle with FS concerns their lack of exposure. It has been proposed that the acquisition of FS essentially relies on extensive exposure (González Fernández and Schmitt 2015). However, although natural language is highly formulaic, individual formulaic sequences do not occur as frequently as single words (e.g. Moon 1998). One way that learners could improve their knowledge of FS is through incidental learning activities. However, only a few studies have examined incidental learning of FS from exposure to L2 input (Choi 2017; Frumuselu et al. 2015; Macis 2018; Pellicer-Sánchez 2017; Peters 2012; Webb, Newton and Chang 2013). It has been shown that FS can be learned incidentally from reading texts.

To our knowledge, only one study has explored incidental learning of FS through viewing (subtitled TV). Frumuselu et al. (2015) conducted a longitudinal study in which L2 learners with various L1s and levels of proficiency watched 13 episodes of an English TV programme with L1 subtitles (captions) or L2 subtitles. The findings showed that it was possible to learn FS from subtitled TV. However, it remains unclear whether FS can be picked up incidentally from watching TV without subtitles or captions. Further research into the incidental learning of FS from audiovisual input thus seems warranted.

Incidental learning: the role of prior knowledge and item-related factors

Incidental learning is a slow and incremental process, which relies heavily on a learner's level of proficiency and prior vocabulary knowledge (Hulstijn 2013). Moreover, since individual words and FS each have their own learning burden (Schmitt 2010), the effectiveness of incidental learning might additionally depend on item-related factors. Some of the factors that have been shown to affect the learning of single words are:

- prior vocabulary knowledge (Montero Perez et al. 2014; Peters and Webb 2018; Peters, Heynen and Puimège 2016): learners with a larger vocabulary size tend to learn more words incidentally.
- corpus frequency (Vidal 2003): frequent words tend to be learned more easily than infrequent words.
- concreteness (Pichette, De Serres and Lafontaine 2012): concreteness is the degree to which a word refers to a perceptible entity (Brysbaert, Warriner and Kuperman 2014). Concreteness has been shown to facilitate processing and learning of words.
- word length (Campoy 2008): short words tend to be learned more easily than long words.
- part of speech (Campoy 2008): nouns tend to be more concrete, imageable and meaningful than verbs (Crossley, Subtirelu and Salsbury 2013), and may also be acquired more easily.

With regard to the learning of FS, phraseological as well as distributional factors have been shown to mediate learning. Some of these are:

- collocate-node relationship (Peters 2016): adjective–noun combinations tend to be learned more easily than other types of FS. Verb-particle combinations such as phrasal verbs are challenging to L2 learners, because they lack semantic transparency and often have multiple meanings (Garnier and Schmitt 2016).
- association strength (Nguyen and Webb 2016): association strength is the strength of co-occurrence of two or more words, based on corpus-derived measures such as mutual information

(MI) and t-score. Mutual information measures the strength of attraction between two words by comparing their observed frequency of co-occurrence to the expected frequency of co-occurrence. An MI score above 0 corresponds to higher co-occurrence than is expected by chance, i.e. attraction between the words, whereas a score below 0 corresponds to lower co-occurrence than is expected by chance, or repulsion (Evert 2008). MI has been shown to assign inflated scores to word combinations that contain low-frequency components (Evert 2008). As a result, FS with high MI scores tend to be more difficult to acquire because they often contain low-frequency words (Nguyen and Webb 2016).

The factors mentioned above have mainly been studied in the context of intentional vocabulary learning (e.g. Peters 2016) or incidental learning from written texts (Pellicer-Sánchez 2017). Not much is known about the effects of item-related factors in incidental vocabulary learning from audiovisual input.

Rationale and research questions

To date, very little is known about the incidental learning of formulaic sequences and the potential of TV viewing for learning FS. There is some evidence now that single words can be learned from watching TV. However, it remains unclear whether FS can be learned to the same extent. As a result, more research into the learning of FS is needed, if we want obtain a fuller picture of the learning process of FS and the factors affecting this process. Given the abundance of FS in everyday spoken language and audiovisual input (Lin 2014), television could be an important source of formulaic language for L2 learners. The aim of this exploratory study is to investigate whether, like single words, FS can be learned from watching a television programme and which item-related factors might affect the learning process. The following research questions were addressed in this study:

(1) Can single words and formulaic sequences be learned from watching L2 television?
(2) Which item-related and learner-related factors affect the incidental learning of single words from watching L2 television?
(3) Which item-related and learner-related factors affect the incidental learning of formulaic sequences from watching L2 television?

To answer the research questions, an exploratory experiment was conducted adopting a pre-test post-test within-subject design. Twenty participants watched an English TV programme without captions or subtitles. English words and formulaic sequences occurring in the programme were tested before and after the treatment. There was no control group, but distractor items were added to the pre- and post-tests in order to control for a test effect.

Method

Participants

Twenty first-year business students at a Flemish university (L1 = Dutch) aged between 19 and 21 (average age: 19.5) took part in the experiment. All participants had received at least five years of instruction in English, as the starting age for English in Flanders is 13–14 years old. Their level of proficiency could be considered B1 to B2 of the Common European Framework of Reference. However, in order to take into consideration individual differences in proficiency, the participants' vocabulary size was measured by means of Nation and Beglar's (2007) Vocabulary Size Test (see Instruments section). Their vocabulary size ranged between 8200 and 15,000 word families (average 11,450). Given the study's focus on audio-visual input, it should be mentioned that learners in the present study

were used to watching English language TV, as was shown in a questionnaire tapping into learners' TV viewing habits.

Materials

Audiovisual input

The audiovisual input selected for this study was a 30-minute excerpt of *Dragon's Den*, a British reality TV programme in which budding entrepreneurs pitch their business ideas to a panel of successful business people. The programme is abundant with business English which makes it an interesting source of vocabulary for business students. By entering the script of the excerpt into the Vocabprofile section of the Compleat lexical tutor (Cobb n.d.), it was found that 87.32% of the words in the input belong to the 1000 most frequent words in the British National Corpus/Corpus for Contemporary American English (BNC/COCA), and 92.85% belonged to the 2000 most frequent words.

Target items

A diverse set of 15 single words and 20 FS was selected from the input (see Table 1 for the single words and Table 2 for the formulaic sequences). Unlike in some reading studies (Pellicer-Sánchez 2017; Webb, Newton and Chang 2013), we did not manipulate the input or target items for reasons of ecological validity. However, because a number of factors have been found to affect the learning of single words and FS, some of these were taken into account in the analyses. The following factors were taken into account in the analyses for the learning of single words:

- corpus frequency
- item length (= number of syllables)
- part of speech
- concreteness

Raw corpus frequencies were obtained from the Spoken and Fiction subcorpora of the COCA (Davies 2008). Concreteness was obtained from Brysbaert, Warriner and Kuperman's (2014) norms. The norms are based on ratings from 4000 participants obtained through internet crowd-sourcing. As opposed to other concreteness norms (e.g. Paivio, Yuille and Madigan 1968), the ratings are available for a large number of lemmas (approx. 40,000), and are based on all modes of perception. The norms can be found on Marc Brybaert's website (http://crr.ugent.be/archives/1330). Because only one word was a Dutch cognate ('empathise'), this factor was not taken into account. Frequency of occurrence in the input was also not taken into account since only three items occurred more than once.

Table 1. Target items – single words.

	Corpus frequency	Frequency of occurrence	Length (syllables)	Part of speech	Concreteness
quirky	1963	2	2	Adjective	2.12
engaging	2568	1	3	Adjective	1.72
fledgling	1785	1	2	Adjective	2.76
articulate	1388	1	4	Adjective	2.31
scalable	216	1	3	Adjective	2.5
to entice	2207	1	2	Verb	1.12
to empathise	800	1	3	Verb	2.23
to forecast	3786	2	2	Verb	2.78
interrogate	1957	1	4	Verb	2.48
quandary	829	1	3	Noun	1.75
longevity	2626	1	4	Noun	1.9
ardor	426	1	2	Noun	2.59
clarification	1714	1	5	Noun	1.93
retail	12200	8	2	Noun	2.92
venture	12314	1	2	Noun	2.6

Table 2. Target items – formulaic sequences.

	Corpus freq.	Freq. of occ.	Length (syll.)	Collocate-node rel.	MI	Type
corporate event	63	1	4	Adj.-noun	3.37	lexical coll.
advance booking	19	1	4	Adj.-noun	8.77	lexical coll.
subliminal message	78	1	6	Adj.-noun	10.53	lexical coll.
raw material	2351	1	5	Adj.-noun	9.91	lexical coll.
vast number	498	2	3	Adj.-noun	5.08	lexical coll.
master × art	366	1	5	Verb-noun	5.53	lexical coll.
capture × imagination	395	1	10	Verb-noun	6.73	lexical coll.
hold × own	1422	1	4	Verb-noun	2.1	idiom
catch × eye	2495	1	3	Verb-noun	3.89	idiom
take × punt	7	2	4	Verb-noun	0.41	lexical coll.
run × course	599	1	7	Verb-noun	0.7	lexical coll.
tap into	2879	1	3	Verb-particle	5.51	gramm. coll.
come about	2830	1	3	Verb-particle	0.54	phrasal verb
head back	3687	1	2	Verb-particle	5.59	phrasal verb
shy away	1472	1	3	Verb-particle	10.63	phrasal verb
turn over	5724	3	3	Verb-particle	3.87	gramm. coll.
drill down	203	1	2	Verb-particle	4.76	phrasal verb
down the line	1888	1	2	Prep.-noun	2.95	idiom
up and running	1287	1	3	Prep.-adj.	1.33	binomial
economies of scale	497	1	6	Noun-noun	5.09	compound

For the learning of FS, the following factors were analysed:

- corpus frequency
- item length (= number of syllables)
- phraseological type
- collocate-node relationship
- mutual information (MI)

The phraseological type of each FS was determined based on Granger and Paquot's (2008) classification. Five types were distinguished: lexical collocations (e.g. '*subliminal message*'), idioms (e.g. '*catch × eye*'), binomials ('*up and running*'), grammatical collocations (e.g. '*turn over*') and phrasal verbs (e.g. '*head back*'). In terms of collocate-node relationship, the following types were included: verb-noun, verb-particle, adjective–noun, noun-noun, preposition-noun, and preposition-adjective. Mutual information (MI) scores ranged between 0.41 and 10.63 (average 4.75). Mutual information tends to favour low-frequency and specialised word combinations (Gablasova, Brezina and McEnery 2017). Because we wanted to include different types of FS, including those containing high-frequency words, we did not adopt the commonly held threshold of 3 (Schmitt 2010), but instead included FS with lower MI scores if these could be considered conventional based on their corpus frequency.

All of the selected target items were incongruent, i.e. did not have a literal word-for-word translation in Dutch. We could not find word-for-word translations for any of the items in online bilingual dictionaries or translation websites, although one item ('*subliminal message*') has a single-word literal translation. Congruency was therefore not included as a variable possibly affecting the learning of FS. Other factors that were not taken into account include adjacency of the component words in the input and grammatical variability. Because only three FS occurred more than once in the input, occurrence frequency was also not included in the analyses. Corpus frequencies of the component words were not included as a separate factor because this would likely cause multi-collinearity with MI and possibly with corpus frequency of the FS as a whole.

Instruments

Vocabulary size test
Because learners' prior vocabulary knowledge has been shown to be an important predictor of incidental vocabulary learning from meaning-focused input (e.g. Peters and Webb 2018), a vocabulary

size test was administered to the participants, viz. Nation and Beglar's (2007) Vocabulary Size Test. The test is a frequency-based multiple choice test, containing 100 items that measure receptive knowledge of the form-meaning link. The score multiplied by 200 gives an estimate of the participant's total vocabulary size (Nation and Beglar 2007).

Vocabulary tests

Because vocabulary learning might take place at various levels of knowledge, three vocabulary tests were used to measure knowledge of form and meaning at different levels of strength (form recognition, meaning recall, form recall). To control for a test effect, five low-frequency single words that did not occur in the input were added to the tests as distractors ('*lunacy*', '*to anneal*', '*to bray*', '*bereft*', '*insipid*').

The first test was a form recognition test which measured the ability to recognise the meaning of a spoken or written form, through a multiple-choice cloze format. Participants were asked to complete an English sentence by selecting a missing word or formulaic sequence out of four options. The distractor options in this test were either non-words (for the single words) or incorrect formulaic sequences, i.e. word combinations that could not be found in the COCA. The distractor options for formulaic sequences were composed by replacing one single-word component with a word assigned to the same part of speech and related to the original word in meaning, in order to avoid deduction of the correct combination based on semantic association. The correct FS could only be recognised based on the form of the collocation as a whole. Below are two examples of sentences included in the form recognition test:

(1) The adverts …… the customer into buying things they don't really want.
       ☐ barlone ☐ delead ☐ entice ☐ insuade
(2) He tried to give the …… that he was a man of the people.
       ☐ astute message ☐ subliminal message ☐ circumspect message ☐ contemplative message

The second test measured knowledge at the level of form recall, i.e. the ability to supply the form of a given meaning. As shown below, participants were asked to write the form of the target words and formulaic sequences based on a Dutch translation or short description. The first letter of each component word was given to avoid elicitation of other potential word combinations.

schaalvoordelen = e _ _ _ _ _ _ _ _ of s _ _ _ _

Given that the target items were presented orally in the input, spelling mistakes were ignored as long as the response was comprehensible (e.g. '*corperate event*', '*sublimenal message*', '*interoggate*'), following Sonbul and Schmitt (2013).

The third test measured meaning recall, i.e. the ability to supply the meaning of a given form. Participants were asked to translate the English words and formulaic sequences into Dutch or describe their meaning. Both the spoken and written form of the target items were given:

economies of scale = …

Questionnaire

In addition to the vocabulary tests, the participants completed a short questionnaire about their comprehension of the gist of the content, their general comprehension, TV viewing habits (e.g. how often they watch English television without subtitles), and familiarity with the television programme. The questionnaire can be found in the Appendix.

Procedure

The data were collected in two sessions. Two weeks before the treatment the vocabulary size test and pre-tests were administered. To minimise a potential test effect, the form recall test was administered first, followed by the form recognition test, the vocabulary size test, and finally the meaning recognition test. The vocabulary size test was completed within 35 minutes, the pre-tests took 15 minutes

each. Two weeks after the pre-tests, participants watched the excerpt from *Dragon's Den*, after which they completed the questionnaire and the immediate post-tests. The post-tests were administered in the same order as the pre-tests. To avoid a test order effect, the items comprising the post-tests appeared in a different order from the items of the pre-tests. Five target items were known by 18 of the participants on the form recall pre-test, and were therefore no longer included in the posttests; '*scalable*', '*interrogate*', '*venture*', '*raw materials*' '*economies of scale*'. After the post-tests, the participants were debriefed about the aims of the study.

Results

Questionnaire

The answers to the questionnaire indicated that the participants had no trouble understanding the gist of the content and that they had prior experience of viewing English television. All participants indicated that they watched English film or television with Dutch subtitles on a daily or a weekly basis. They also watched English video without subtitles, albeit less often (once a week/month). All participants were familiar with the television programme *Dragon's Den* and had watched at least one episode before the treatment. However, none of the participants had watched the episode used in the treatment.

Vocabulary tests[1]

The vocabulary tests were scored dichotomously; a correct answer received a score of 1, an incorrect answer received a score of 0. The descriptive results are reported in Tables 3 and 4.

Learning gains were found for all levels of vocabulary knowledge. Although scores on the form recall test were lower than on the meaning recall and meaning recognition tests, absolute learning gains (i.e. the difference between the pre- and post-test scores) were the highest for this area of knowledge. Learning gains were higher for formulaic sequences than for single words, but some learning gains were also made for the distractors, possibly due to a test effect. The 95% confidence intervals indicate that differences between the pre- and post-tests were nevertheless larger for the target items than for the distractors. Closer examination of the individual results of the form recognition test revealed that many scores per item were higher on the pre-test than on the post-test, which may be due to guessing. Guessing effects are typical of multiple-choice tests and often lead to overestimation of scores (Gyllstad, Vilkaitė and Schmitt 2015). Therefore, the results of the form recognition test were not further analysed.

Research question 1: Can single words and formulaic sequences be learned from watching L2 television?

Logistic regression analyses with repeated measures (Generalised Estimating Equations (GEE) in SPSS) were performed on the scores of each test to find out if there were significantly more learning gains for target items than for distractors and if the treatment (watching TV) had an effect. The dependent

Table 3. Learning gains – target items and distractors.

	Pre-test M (SD)	95% CI	Post-test M (SD)	95% CI	Abs. gains M (SD)	95% CI
Target items (max = 35)						
Form recognition	20.45 (4.97)	18.12–22.78	23.7 (3.7)	21.97–25.43	3.45 (2.78)	2.15–4.75
Form recall	3.35 (2.97)	1.96–4.74	9.3 (5.67)	6.65–11.95	5.95 (3.73)	4.20–7.70
Meaning recall	12.15 (6.16)	9.27–15.03	15.6 (5.71)	12.93–18.27	3.6 (2.44)	2.46–4.74
Distractors (max = 5)						
Form recognition	1.65 (1.09)	1.14–2.16	2.25 (1.29)	1.65–2.85	0.6 (1.57)	−0.13–1.33
Form recall	0.05 (0.22)	0.39–1.31	0.4 (0.5)	0.93–2.17	0.35 (0.49)	0.36–1.04
Meaning recall	0.85 (0.99)	−0.05–0.15	1.55 (1.32)	0.16–0.64	0.7 (0.73)	0.12–0.58

Table 4. Learning gains – single words and formulaic sequences.

	Pre-test M (SD)	95% CI	Post-test M (SD)	95% CI	Abs. gains M (SD)	95% CI
Single words (max = 15)						
Form recognition	7.80 (2.24)	6.75–8.85	8.85 (1.84)	7.99–9.71	1.05 (2.35)	−0.05–2.15
Form recall	1.25 (1.16)	0.71–1.79	3.25 (2.34)	2.16–4.34	2 (1.59)	1.26–2.74
Meaning recall	4.20 (2.26)	3.14–5.26	5.52 (2.29)	4.18–6.32	1.05 (1.54)	0.33–1.77
FS (max = 20)						
Form recognition	12.65 (3.42)	11.05–14.25	14.85 (2.23)	13.81–15.89	2.20 (2.26)	1.14–3.26
Form recall	2.1 (2.13)	1.11–3.09	6.05 (3.43)	4.45–7.65	3.95 (2.48)	2.79–5.11
Meaning recall	7.95 (4.25)	5.96–9.94	10.35 (3.91)	8.52–12.18	2.4 (2.06)	1.43–3.37

variable in the models (one model for the meaning recall test, and one for the form recall test) was the score on the post-test given that the score on the pre-test equalled 0. Whether or not a test item was a distractor was entered as a dichotomous factor in the analyses. The participants' vocabulary size was entered as covariate.

On the form recall test, 'type of item (= target or distractor)' was a significant predictor of learning ($p < .001$, $B = 1.429$, $Exp(B) = 4.174$), as was vocabulary size ($p < .001$, $B = 0.069$, $Exp(B) = 1.071$). This indicates that target items were learned significantly better than distractors. Learning gains could therefore be ascribed to the treatment, and not to the vocabulary tests.

On the meaning recall test, the factor 'type of item' did not significantly predict learning ($p = .24$), while vocabulary size did ($p = .036$, $B = 0.035$, $Exp(B) = 1.036$). This indicates that both target items and distractors were learned to the same extent on the meaning recall test. For that reason, the effects of item-related and learner-related factors were only analysed for the form recall test results.

Research question 2: Which item-related and learner-related factors affected learning of single words?

Logistic regression analyses with repeated measures were performed on the form recall test results in order to investigate the relationship between vocabulary learning and item-related factors. Two models were designed: one for single words, one for formulaic sequences. We checked for the assumptions of logistic regressions (multi-collinearity, linearity of independent variables and log odds, ratio of the number of observations and the number of variables).

For the single words, the following variables were entered into the regression model: corpus frequency (logarithmically transformed), length in syllables, part of speech, concreteness. Through a backward stepwise selection, the non-significant predictors were removed one by one, resulting in the final models including only predictors with a p-value lower than .05. The results are presented in Table 5. Four parameters were significantly related to learning: corpus frequency, item length, concreteness, and learners' vocabulary size. Part of speech did not significantly predict learning. Concreteness was the most important predictor of learning (see $Exp(B)$ in Table 5); the odds of a correct response on the form recall post-test were 5.03 times higher if the concreteness level increased by 1 unit. Vocabulary size was also positively correlated with learning. The odds of a correct response were 1.12 times higher if a participant's score on the vocabulary size test increased by one. This means that the odds of a correct response were 3.19 times higher if a participant's score on the vocabulary size test increased by ten ($Exp(10*B)$), i.e. if their vocabulary size increased by 2000 word families. Large differences in vocabulary size could thus account for differences in learning gains. The effects of corpus frequency and length were also positively correlated with learning.

Research question 3: Which item-related and learner-related factors affected learning of formulaic sequences?

For the learning of formulaic sequences, the following covariates and factors were included in the logistic regression: corpus frequency, length in syllables (as appearing in the input), collocate-node

Table 5. Generalised estimating equations (GEE) for single words.

	Form recall		
	B	*p*	*Exp* (*B*)
Vocabulary size	0.12	0	1.12
Length	0.56	0	1.76
Concreteness	1.62	<0	5.03
Corpus frequency	1.39	<0	4.00
PoS		ns	

relationship, mutual information, phraseological type. Again through a backward stepwise selection, the non-significant predictors were removed one by one, resulting in the final models including only predictors with a *p*-value lower than .05. Because phraseological type and collocate-node relationship were highly correlated (>.70), only collocate-node relationship was included in the analyses given its higher correlation with the posttest scores.

All factors were significantly related to learning in the form recall test (see Table 6). The most important predictor of learning was the item's collocate-node relationship. The odds of a correct response were 30.64 times higher if the FS was an adjective–noun combination. The effect size was also very large for verb-noun combinations (12.05), indicating that verb-particle combinations (phrasal verbs and grammatical collocations) had much smaller learning gains than the other two types. For mutual information, a negative correlation was found. The stronger the association between the component words, the smaller the learning gains. Similarly, a negative relationship was found between item length and learning. Specifically, the odds of a correct response were 1.25 (1/*Exp*(*B*)) times higher if the FS was one syllable shorter. As was the case for the single words, learners' vocabulary size was positively correlated with learning FS. The odds of a correct response were 2.18 times higher if the participant's score on the vocabulary size test increased by ten, i.e. if their vocabulary size increased by 2000 word families.

Discussion

Research question 1: Can single words and formulaic sequences be learned incidentally from L2 television?

Our findings tentatively suggest that both single words and formulaic sequences can be learned incidentally through TV viewing in terms of form recall. On the form recall test, learning gains were significantly higher for target items than for distractors. We can therefore assume that the items were learned from the treatment. Even though previous research has shown that incidental vocabulary learning is a slow and incremental process (Nation 2001), and that one encounter with a word or FS is unlikely to result in a strong form-meaning link (Hulstijn 2013), our findings suggest that different word knowledge aspects can be learned even from a single exposure. Previous research into the effects of TV viewing has mainly focused on learning the meaning of target items. Our study adds that learning might also occur at the level of form recall, a knowledge

Table 6. Generalised estimating equations (GEE) for formulaic sequences.

	Form recall		
	B	*p*	*Exp.* (*B*)
Vocabulary size	0.08	<0.001	1.08
Coll.-node rel. = adjective-noun	3.42	<0.001	30.64
Coll.-node rel. = verb-noun	2.49	<0.001	12.05
Length	−0.23	0.04	0.80
MI	−0.16	0.001	0.85
Corpus freq.	0.55	0.016	1.74

aspect that is generally considered to be more difficult than meaning recall and recognition (Nation 2001).

On the meaning recall test, however, no significant difference was found between the learning gains for target items and distractors. The lack of this significant difference might be explained by a test effect. Because the form recall test, in which the target items' definitions were provided, was administered before the meaning recall test, it is possible that participants associated some of the target items' forms on the meaning recall test with definitions they remembered from the form recall test. A second explanation might be the low number of participants, which resulted in considerable variation in learning gains. Finally, the low number of target items and distractor items in particular might also have played a role. Consequently, no conclusions can be drawn for the effect of TV viewing at the level of meaning recall.

Research question 2: Which item-related and learner-related factors affect the incidental learning of single words from watching L2 television?

In answer to the second research question, the results of the logistic regression analysis show that item-related variables as well as the participants' vocabulary size played a role in the learning of single words. Our study adds to the growing body of evidence that learners' vocabulary size is positively correlated with learning. The more words learners know, the more likely they are to pick up new words incidentally from viewing. This is in line with previous findings that prior vocabulary knowledge predicts incidental learning (Montero Perez et al. 2014; Peters, Heynen and Puimège 2016; Peters and Webb 2018).

Secondly, a positive correlation was also found for concreteness. More concrete target words (e.g. 'empathise', 'quirky') were learned better than more abstract words (e.g. 'engaging', 'to entice'), which supports findings of previous studies that concrete words are easier to learn incidentally (Pichette, De Serres and Lafontaine 2012).

Thirdly, corpus frequency was positively correlated with learning. In psycholinguistic research, corpus frequency is considered a good proxy for word familiarity (Kuperman and Van Dyke 2013). Because words that occur frequently in a language tend to be encountered more often, they are also more likely to be acquired. It is possible that the participants in our study already had partial knowledge of the most frequent items (e.g. 'quirky', 'engaging') before the experiment, which they strengthened at the level of form recall. The correlation found for corpus frequency might thus to some extent reflect familiarity with the target items.

Finally, there was a small positive correlation between learning and word length. Longer target words may have been more salient in the aural input and therefore more likely to be learned. However, some of the longer items such as 'articulate' and 'quandary' were not learned well, indicating that the role of length was less clear than that of concreteness and vocabulary size.

Research question 3: Which item-related and learner-related factors affect the incidental learning of formulaic sequences from watching L2 television?

All item-related factors vocabulary size played a role in the incidental learning of FS. First, a positive correlation was found between the participants' vocabulary size and learning, which suggests that learners' prior vocabulary knowledge may facilitate the incidental learning of FS from viewing. The facilitating role of prior vocabulary knowledge on the acquisition of FS has also been found in one previous study (Peters 2016).

Secondly, collocate-node relationship was an important predictor of learning. In line with previous findings (Peters 2016), the highest learning gains were found for adjective–noun combinations, whereas the lowest learning gains were found for verb-particle combinations. As has been mentioned in previous studies (Garnier and Schmitt 2016), phrasal verbs often have opaque or idiomatic meanings (e.g. 'drill down', 'come about'). The verb-particle combinations in our study may therefore

have been semantically less accessible than other types of FS. Moreover, the results of the form recall test show that the participants were sometimes able to provide the particle, but not the verb. Verbs such as 'head' in 'head back' and 'shy' in 'shy away' do not often occur in other word combinations, and may have been unfamiliar to the participants.

Thirdly, a negative correlation was found between length and the learning of FS. This is surprising, given that the longest FS were adjective–noun combinations and some of the shortest FS were verb-particle combinations. The effect of item length may be due to a few outliers that were not learned well, such as 'subliminal message', and to the limited number of target items in general, making it difficult to draw any strong conclusions.

Another factor that negatively predicted learning was MI. In their study on intentional learning of collocations Nguyen and Webb found (2016) a negative correlation between mutual information (MI) and learning, which they ascribed to the low-frequency bias of MI, because FS that consist of low-frequency words tend to have higher MI scores. In our study too, FS with low-frequency components had high MI scores and were not learned well, possibly because of participants' lack of knowledge of the component words (e.g. 'shy away', 'subliminal message').

Finally, a small correlation was found for corpus frequency. It has been suggested that corpus frequency does not predict L2 knowledge of FS to the same extent as it predicts single word knowledge (González Fernández and Schmitt 2015), which might explain the small correlation found in our study. However, given the small sample size and the possible effect of prior knowledge, no strong conclusions should be drawn.

Limitations and suggestions for future research

The study reported in this article is based on a small-scale experiment and has a number of limitations. Due to the small sample of participants, the generalisability of our findings is limited. It should also be noted that the participants in our experiment are all intermediate to high-proficiency EFL learners who are used to watching English television in their spare time. Therefore, the findings of this study cannot be generalised to other EFL contexts. Another limitation concerns the limited number of target items. A larger sample of items might decrease the chance of a test effect. Furthermore, the findings of our study might underestimate the amount of learning from the input because the results of the form recognition and meaning recall test were not interpreted. The results of the form recognition test were unreliable due to a guessing effect, which has also been found in other studies for this test format (Peters and Webb 2018). The results of the meaning recall test may have been distorted due to a test effect. Nevertheless, it is likely that some vocabulary was learned at these two levels of knowledge. Future studies should include a larger sample of participants and target items, in order to be better able to generalise the findings. As previous research has shown, formulaic language is a vast and complex phenomenon that cannot be represented by a sample of 15 FS. Finally, given that formulaic language is not all the same, more fine-grained analyses are needed to further investigate the role of influencing factors such as collocate-node relationship and strength of association.

Conclusions

The aim of this study was to explore the incidental learning of single words and formulaic sequences through TV viewing. The results of our experiment show that words and formulaic sequences can be learned from exposure to L2 television, even at the level of form recall. Learning gains were affected by prior vocabulary knowledge and item-related factors such as concreteness and collocate-node relationship. Given that formulaic language is ubiquitous in spoken discourse and on television, watching L2 television might be an effective way for learners to expand their knowledge of formulaic sequences outside the classroom. More research is needed to explore the potential of L2 television for incidental learning of FS, and to improve our understanding of the factors affecting the learning process of single words and formulaic language.

Note

1. The Kolmogorov-Smirnov test for normality showed that only the scores on the form recall post-test were normally distributed ($p < .146$). All other data deviated from normality (form recall pretest: $p < .001$, meaning recall pretest: $p = .014$, meaning recall posttest: $p = .041$).

Disclosure statement

No potential conflict of interest was reported by the authors.

References

Bahns, J. and M. Eldaw. 1993. Should we teach EFL students collocations? *System* 21: 101–14. doi:10.1016/0346-251X(93)90010-E.

Brysbaert, M., A.B. Warriner and V. Kuperman. 2014. Concreteness ratings for 40 thousand generally known English word lemmas. *Behavior* 46, no. 3: 904–11. doi:10.3758/s13428-013-0403-5.

Campoy, G. 2008. The effect of word length in short-term memory: is rehearsal necessary? *Quarterly Journal of Experimental Psychology* 61, no. 5: 724–34. doi:10.1080/17470210701402364.

Choi, S. 2017. Processing and learning of enhanced English collocations: an eye movement study. *Language Teaching Research* 21, no. 3: 403–26. doi:10.1177/1362168816653271.

Cobb, T. n.d. *Compleat Lexical Tutor* (Version 8.3) [Online resource]. http://www.lextutor.ca/.

Crossley, S.A., N. Subtirelu and T. Salsbury. 2013. Frequency effects or context effects in second language word learning. *Studies in Second Language Acquisition* 35: 727–55. doi:10.1017/S0272263113000375.

Davies, M. 2008. The Corpus of Contemporary American English (COCA): 560 million words 1990-present. https://corpus.byu.edu/coca/.

Erman, B. and B. Warren. 2000. The idiom principle and the open choice principle. *Text - Interdisciplinary Journal for the Study of Discourse* 20, no. 1: 29–62. doi:10.1515/text.1.2000.20.1.29.

European Union. 2018. *Standard Eurobarometer 88 report: media use in the European Union*. doi:10.2775/116707.

Evert, S. 2008. Corpora and collocations. Extended manuscript of chapter 58. In *Corpus Linguistics. An International Handbook*, ed. A. Lüdeling and M. Kytö, 1212–1248. Berlin: Mouton de Gruyter.

Frumuselu, A.D., S. De Maeyer, V. Donche and M.d.M.G. Colon Plana. 2015. Television series inside the EFL classroom: bridging the gap between teaching and learning informal language through subtitles. *Linguistics and Education* 32: 107–17. doi:10.1016/j.linged.2015.10.001.

Gablasova, D., V. Brezina and T. McEnery. 2017. Exploring learner language through corpora: comparing and interpreting corpus frequency information. *Language Learning* 67: 130–54.

Garnier, M. and N. Schmitt. 2016. Picking up polysemous phrasal verbs: how many do learners know and what facilitates this knowledge? *System* 59: 29–44. doi:10.1016/j.system.2016.04.004.

González Fernández, B. and N. Schmitt. 2015. How much collocation knowledge do L2 learners have? The effects of frequency and amount of exposure. *ITL - International Journal of Applied Linguistics* 166, no. 1: 94–126. doi:10.1075/itl.166.1.03fer.

Granger, S. 1998. Prefabricated patterns in advanced EFL writing: collocations and formulae. In *Phraseology: Theory, Analysis and Applications*, ed. A.P. Cowie, 79–100. Oxford: Oxford University Press.

Granger, S. and M. Paquot. 2008. Disentangling the phraseological web. In *Phraseology: An Interdisciplinary Perspective*, ed. S. Granger and F. Meunier, 27–49. Amsterdam and Philadelphia: Benjamins.

Gyllstad, H., L. Vilkaitė and N. Schmitt. 2015. Assessing vocabulary size through multiple-choice formats: issues with guessing and sampling rates. *ITL - International Journal of Applied Linguistics* 166, no. 2: 278–306. doi:10.1075/itl.166.2.04gyl.

Horst, M., T. Cobb and P. Meara. 1998. Beyond *A Clockwork Orange*: acquiring second language vocabulary through reading. *Reading in a Foreign Language* 11: 207–23.

Hulstijn, J.H. 2013. Incidental learning in second language acquisition. In *The Encyclopedia of Applied Linguistics*, ed. C.A. Chapelle, 2632–40. Chichester: Wiley-Blackwell. doi:10.1002/9781405198431.wbeal0530.

Kuperman, V. and J.A. Van Dyke. 2013. Reassessing word frequency as a determinant of word recognition for skilled and unskilled readers. *Journal of Experimental Psychology and Human Perception and Performance* 39, no. 3: 802–23. doi:10.1037/a0030859.

Larsen-Freeman, D. 1983. The importance of input in second language acquisition. In *Pidginization and Creolization as Second Language Acquisition*, ed. R. Anderson, 87–93. Rowley, MA: Newbury House.

Laufer, B. and T. Waldman. 2011. Verb-noun collocations in second language writing: a corpus analysis of learners' English. *Language Learning* 61, no. 2: 647–72. doi:10.1111/j.1467-9922.2010.00621.x.

Lin, P.M.S. 2014. Investigating the validity of internet television as a resource for acquiring L2 formulaic sequences. *System* 42: 164–76.

Macis, M. 2018. Incidental learning of duplex collocations from reading: three case studies. *Reading in a Foreign Language* 30, no. 1: 48–75.

Montero Perez, M., E. Peters, G. Clarebout and P. Desmet. 2014. Effects of captioning on video comprehension and incidental vocabulary learning. *Language Learning and Technology* 18, no. 1: 118–41.

Moon, R.E. 1998. *Fixed Expressions and Text: A Study of the Distribution and Textual Behaviour of Fixed Expressions in English*. Oxford: Clarendon.

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

Nation, I.S.P. 2006. How large a vocabulary is needed for reading and listening? *The Canadian Modern Language Review / La Revue Canadienne Des Langues Vivantes* 63, no. 1: 59–82. doi:10.1353/cml.2006.0049.

Nation, I.S.P. and D. Beglar. 2007. A vocabulary size test. *The Language Teacher* 31: 9–13.

Nesselhauf, N. 2003. The use of collocations by advanced learners of English and some implications for teaching. *Applied Linguistics* 24: 223–42. doi:10.1093/applin/24.2.223.

Nesselhauf, N. 2005. *Collocations in a Learner Corpus.* Amsterdam: John Benjamins. doi:10.1075/scl.14.

Neuman, S.B. and P. Koskinen. 1992. Captioned television as comprehensible input: effects of incidental word learning from context for language minority students. *Reading Research Quarterly* 27, no. 1: 94–106. doi:10.2307/747835.

Nguyen, T.M.H. and S. Webb. 2016. Examining second language receptive knowledge of collocation and factors that affect learning. *Language Teaching Research* 21, no. 3: 298–320. doi:10.1177/1362168816639619.

Paivio, A., J.C. Yuille and S.A. Madigan. 1968. Concreteness, imagery, and meaningfulness values for 925 nouns. *Journal of Experimental Psychology* 76: 1–25.

Pellicer-Sánchez, A. 2017. Learning L2 collocations incidentally from reading. *Language Teaching Research* 21, no. 3: 381–402. doi:10.1177/1362168815618428.

Pellicer-Sánchez, A. and N. Schmitt. 2010. Incidental vocabulary acquisition from an authentic novel: do things fall apart? *Reading in a Foreign Language* 22, no. 1: 31–55.

Peters, E. 2012. Learning German formulaic sequences: the effect of two attention-drawing techniques. *The Language Learning Journal* 40, no. 1: 65–79. doi:10.1080/09571736.2012.658224.

Peters, E. 2016. The learning burden of collocations: the role of interlexical and intralexical factors. *Language Teaching Research* 20, no. 1: 113–38. doi:10.1177/1362168814568131.

Peters, E. 2018. The effect of out-of-class exposure to English language media on learners' vocabulary knowledge. *ITL – International Journal of Applied Linguistics* 169, no. 1: 142–68.

Peters, E. and S. Webb. 2018. Incidental vocabulary acquisition through viewing L2 television and factors that affect learning. *Studies in Second Language Acquisition* 40: 551–77. 1–27. doi:10.1017/S0272263117000407.

Peters, E., E. Heynen and E. Puimège. 2016. Learning vocabulary through audiovisual input: The differential effect of L1 subtitles and captions. *System* 63: 134–48. doi:10.1016/j.system.2016.10.002.

Peters, E., A. Noreillie, K. Heylen, B. Bulté and P. Desmet. 2019. The impact of instruction and out-of-school exposure to foreign language input on learners' vocabulary knowledge in two languages. *Language Learning*. doi:10.1111/lang.12351.

Pichette, F., L. De Serres and M. Lafontaine. 2012. Sentence reading and writing for second language vocabulary acquisition. *Applied Linguistics* 33, no. 1: 66–82. doi:10.1093/applin/amr037.

Saragi, T., I.S.P. Nation and F. Meister. 1978. Vocabulary learning and reading. *System* 6: 72–8.

Schmitt, N. 2010. *Researching Vocabulary: A Vocabulary Research Manual*. London: Palgrave Macmillan.

Sonbul, S. and N. Schmitt. 2013. Explicit and implicit lexical knowledge: acquisition of collocations under different input conditions. *Language Learning* 63, no. 1: 121–59.

Vidal, K. 2003. Academic listening: a source of vocabulary acquisition? *Applied Linguistics* 24, no. 1: 56–89. doi:10.1093/applin/24.1.56.

Webb, S. and M.P.H. Rodgers. 2009. Vocabulary demands of television programmes. *Language Learning* 59, no. 2: 335–66. doi:10.1111/j.1467-9922.2009.00509.x.

Webb, S., J. Newton and A. Chang. 2013. Incidental learning of collocation. *Language Learning* 63, no. 1: 91–120. doi:10.1111/j.1467-9922.2012.00729.x.

Appendix

Questionnaire on learners' comprehension of the TV programme, TV viewing habits and familiarity with the TV programme.

(1) *Wat vond je van het videofragment?* What was your reaction to the video clip?
    saai (*boring*) 0 - 0 - 0 - 0 - 0 interessant (*interesting*)
    moeilijk te begrijpen 0 - 0 - 0 - 0 - 0 gemakkelijk te begrijpen
    (*difficult to understand*) (*easy to understand*)
(2) Welke 'pitch' (British DJ and MC Academy / Love Da Pop) vond je de beste? Leg kort uit waarom. *Which pitch did you think was the best? Explain briefly why.*

(3) Zou je, als je het geld had, zelf investeren in (één van) de ondernemingen? Leg kort uit waarom wel/niet. *If you yourself had the money, would you invest in any of the propositions? Explain briefly why/why not.*

(4) Heb je iets geleerd uit het fragment? Zo ja, wat heb je geleerd?
Did you learn anything from the clip? If so, what?

(5) Hoe vaak (ongeveer) kijk je naar Engelstalige tv-programmema's, series of films?
nooit / jaarlijks / maandelijks / wekelijks / dagelijks
Approximately how often do you watch English language TV programmes, series or films?
never/once a year/once a month/once a week/ once a day

(6) Hoe vaak (ongeveer) kijk je naar Engelstalige tv-programma's, series of films zonder ondertiteling?
nooit / jaarlijks / maandelijks / wekelijks / dagelijks
Approximately how often do you watch English language TV programmes, series or films without subtitles?
never/once a year/once a month/once a week/ once a day

(7) Kende je het programmema 'Dragon's Den' al? Ja / nee (als het antwoord 'nee' is, ga naar vraag 10)
Were you already familiar with the programme 'Dragon's Den'? Yes/no (If no, go to Question 10)

(8) Hoeveel afleveringen had je hiervoor al bekeken?
(bijna) alle afleveringen / meerdere seizoenen / één seizoen / een paar afleveringen / één aflevering / een stukje van een aflevering
How many episodes have you watched so far?
(almost) all episodes / several series / one series / a couple of episodes / one episode / part of an episode

(9) Had je deze aflevering eerder al bekeken? Ja / nee
Had you seen this episode before? Yes/no

(10) Zou je opnieuw naar 'Dragon's Den' kijken? Waarom wel/niet?
Would you watch 'Dragon's Den' again? Why? Why not?