3-23-2018

The Effect of Content Retelling on Vocabulary Uptake from a TED Talk

Chi-Duc Nguyen
University of Languages and International Studies

Frank Boers
fboers@uwo.ca

The Effect of Content Retelling on Vocabulary Uptake from a TED Talk

Chi-Duc Nguyen

&

Frank Boers

Abstract

This study investigates the potential benefits for incidental vocabulary acquisition of implementing a particular sequence of input-output-input activities. More specifically, EFL learners (n = 32) were asked to watch a TED Talks video, orally sum up its content in English, and then watch the video once more. A comparison group (n = 32) also watched the TED Talks video twice but were not required to sum it up in between. Immediate and delayed post-tests showed significantly better word-meaning recall in the former condition. An analysis of the oral summaries showed that it was especially words which learners attempted to use that stood a good chance of being recalled later. These findings are interpreted with reference to Swain's (e.g., 1995) Output Hypothesis, Laufer and Hulstijn's (2001) Involvement Load Hypothesis, and Nation and Webb's (2011) Technique Feature Analysis. What makes the text-based output task in this experiment fundamentally different from many previous studies which have investigated the merits of text-based output activities is that it was at no point stipulated for the participants that they should use particular words from the input text. The study also illustrates the potential of TED Talks as a source of authentic audio-visual input in EFL classrooms.

It is indisputable that vocabulary knowledge is a hallmark of proficiency and should figure high on language learners' and teachers' agendas (e.g., Stæhr, 2008, and Qian, 2002, for illustrations of the importance of vocabulary knowledge for listening and reading comprehension, respectively). At the same time, there is a consensus that only a certain proportion of available class time can be reserved for deliberate, language-focused learning and teaching (Nation, 2007). After all, sufficient time needs to be devoted to message-focused activities that foster communicative competence. Fortunately, vocabulary can also be acquired incidentally in the course of such message-focused activities, for instance as a by-product of activities where learners are primarily engaged with the content of a text. For example, incidental acquisition can occur when learners read a story, encounter an unfamiliar word that appears crucial to understanding the story, and attempt to figure out its meaning on the basis of contextual cues. In that case, the attention given to the new word serves the purpose of text comprehension rather than being driven by the intention of adding the new word to one's lexical resources. This does of course not preclude the possibility that learners also read texts with a view to expanding their vocabulary knowledge, and so the distinction between incidental and intentional learning is not totally straightforward. A situation where it is probably safe to say that new vocabulary knowledge is the result of intentional vocabulary study instead of incidental acquisition is when learners are told beforehand that a vocabulary test will follow the given activity (Hulstijn, 2001). The present article reports an experimental study where learners were prompted to engage with the content of input materials and were not forewarned about a vocabulary test. The learning conditions investigated here are therefore situated in the realm of incidental vocabulary acquisition as characterized above.

The primary question addressed in this study is whether learners' vocabulary uptake from audio-visual input (more specifically a TED Talks video) is positively influenced by asking them to sum up its content before engaging with the same input text once more.

The effect of text-based output activities on vocabulary uptake has been explored in a number of studies before (see further below), but, as we shall see, the present investigation is different in at least four ways. First, unlike previous studies, it was not stipulated to the learners which words from the input material they should focus on in the output activity. Instructions to

work specifically with pre-selected words from an input text create a shift from a content-focused activity to a language-focused exercise (and thus from incidental acquisition to intentional study practice). While this has been shown to enhance learners' retention of selected words, it does not address the aforementioned concern over the allocation of class time. Neither does it help teachers to estimate the rate of vocabulary learning that can be expected from text-based communicative activities where students themselves determine what vocabulary to recycle from the text. Second, we used an authentic, unmodified, input text. The texts used in most previous studies on this subject were created or modified for the purpose of drawing learners' attention to particular target words and their meanings, typically through typographic enhancement (e.g., underlining of the words in the text) and through the provision of glossaries. Third, unlike previous research (which often used a mere input-output sequence in their treatment conditions), the present study applied an input-output-input cycle where the learners revisited the input text and could thus compare their own rendering of the content to the original 'model'. Fourth, instead of using a reading text as was done in previous studies on the subject, the present study used audio-visual material—a TED Talk—as the input text. It is not the intention of the present investigation to compare vocabulary uptake from texts presented in different modalities. Rather, the decision to use audio-visual materials was motivated by their increased availability in L2 instructional settings (e.g., King, 2002). While much research has been concerned with vocabulary acquisition from reading (and to a lesser extent from listening), the growing popularity of audio-visual materials calls for more investigations of their affordances for vocabulary acquisition. Some studies have already demonstrated the potential of authentic audio-visual materials for this purpose (e.g., Rodgers & Webb, 2011; Webb, 2015), and there are grounds for believing that audio-visual input offers affordances that are absent from audio input alone (Vanderplank, 2010).

Literature review

Vocabulary acquisition from textual input

Many studies have focused on reading input as a source of incidental vocabulary acquisition (e.g., Horst, Cobb, & Meara, 1998; Pigada & Schmitt, 2006; Waring & Takaki,

2003) and have investigated such variables as repeated encounters with the same words (Chen & Truscott, 2010; Eckerth & Tavakoli, 2012; Horst, Cobb, & Meara, 1998, Laufer & Rozovski-Roitblat, 2015; Rott, 1999; Webb, 2007), the presence of contextual cues (Webb, 2008) and learners' familiarity with and interest in the topic of the text (Lee & Pulido, 2017; Pulido, 2007). While nobody disputes that reading can foster vocabulary knowledge, it is also recognized that vocabulary uptake from content-focused reading alone tends to be a relatively slow process (Laufer, 2003). Ways of modifying texts have been proposed with a view to accelerating and facilitating vocabulary uptake from reading. These include increasing the visual salience of target words by means of typographic enhancement and/or by providing glosses or annotations (e.g., Chun & Plass, 1996; Hulstijn, Hollander, & Greidanus, 1996; Jung, 2016; Ko, 2012; Kost, Fost, & Lenzini, 1999). In addition, the text may be accompanied by comprehension questions that are designed such that learners will need to work out or look up the meaning of particular lexical items in the text to be able to answer the questions (e.g., Hu & Nassaji, 2016; Hulstijn & Laufer, 2001; Peters, 2012; Peters, Hulstijn, Sercu, & Lutjeharms, 2009).

Vocabulary uptake from audio recordings has attracted less attention from researchers (but see, e.g., van Zeeland & Schmitt, 2012, 2013, and Vidal, 2003). Studies which have compared the rates of vocabulary uptake from reading and listening input have shown poorer rates under the latter mode (Brown, Waring, & Donkaewbua, 2008; Vidal, 2011). This is most probably due to the real-time nature of listening, where, unlike in the case of reading, one cannot return to earlier passages to verify one's interpretation (Rost, 2002; Vandergrift, 2007). It may also be due to the challenge that especially lower-proficiency learners experience with parsing the stream of speech into distinct lexical items (Goh, 2000). Listening to an audio recording while reading its transcript (or reading a story while listening to an audio recording of the story) (e.g., Webb & Chang, 2015) can help learners in the latter regard. Another suggestion for assisting learners' listening comprehension is the use of audio-visual materials, because the visual component of video input can offer clues to what is being talked about (e.g., Baltova, 1994; Neuman & Koskinen, 1992). As already mentioned, it is this option of using audio-visual input that was taken in the experiment we report here.

The addition of productive activities

Many studies—often using Laufer and Hulstijn's (2001) *Involvement Load* framework (see further below)—have investigated the benefits of adding vocabulary-focused tasks or exercises to a reading comprehension activity. Table 1 provides an overview (not intended to be exhaustive) of such studies with a brief description of the types of vocabulary tasks or exercises done by the learners in the reading + vocabulary-focus conditions. What most of the studies have in common is that target words are highlighted in the texts and are accompanied by glosses which clarify their meaning. The learners are subsequently asked to further engage with these words, for example, by matching the words to their definitions, translating the words into L1, choosing among from the set of target words to complete gapped sentences, and inventing a new sentence with each of the target words. In some of the studies, the learners were given a post-reading composition task, but almost invariably with the explicit instruction to incorporate the target words and thus also turning these pre-selected words into objects of language study.

Table 1: Text-based 'output' activities used in previous studies with a vocabulary focus

Studies	**Types of activity**
Paribakht & Wesche (1997)	Eight consecutive exercises, ranging from finding and interpreting stipulated words in the text to matching words to their definitions and making sentences with the words
Joe (1998)	Retelling a text with prior instruction to incorporate and clarify unfamiliar words from the text
Hulstijn & Laufer (2001)	a. Selecting words to fill gaps in the text b. Composition writing incorporating stipulated words
Keating (2008)	a. Selecting words to fill gaps in the text b. Sentence writing incorporating stipulated words
Min (2008)	a. Matching words with their definitions b. Selecting words to fill gaps in sentences c. Translation (L1 to L2)

Kim (2008)	a. Selecting words to fill gaps in the text b. Sentence writing incorporating stipulated words c. Composition writing incorporating stipulated words
Peters, et al. (2009)	Providing the L1 translation or an L2 synonym of selected words
Laufer & Rozovski-Roitblat (2011)	a. Matching words with their definitions (or similar) b. Selecting words to fill gaps in sentences c. Translation (L1 to L2)
Peters (2012)	Providing the L1 translation and choosing the correct L2 definition for selected words
Eckerth & Tavakoli (2012)	a. Selecting words to fill gaps in the text b. Composition writing incorporating stipulated words
Lu (2013)	a. Selecting words to fill gaps in a text or in sentences (in one condition three times per word) b. Summary writing incorporating the stipulated words
Laufer & Rozovski-Roitblat (2014)	a. Matching words with their definitions (or similar) b. Selecting words to fill gaps in sentences c. Translation (L1 to L2)
Hu & Nassaji (2016)	a. Choosing the correct definition of target words b. Selecting words to fill gaps in the text
Rassaei (2017)	Various writing tasks (e.g., inventing the ending of a story) incorporating stipulated words
Zou (2017)	a. Selecting words to fill gaps in the text b. Sentence writing incorporating stipulated words c. Composition writing incorporating stipulated words
Yang, et al. (2017)	a. Selecting words to fill gaps in a summary of the text b. Sentence writing incorporating stipulated words c. Writing a short essay related to the theme of the text

While this body of research demonstrates that such extra work done with target words enhances the likelihood that these words will be retained in memory, it also adds to the amount of class time allocated to language-focused learning. This arguably risks compromising efforts to reserve sufficient time for communicative skills development in keeping with, for example, Nation's (2007) recommendations for program design. A potential alternative worth exploring is to try text-based output activities that remain content-focused while at the same time creating opportunities for incidental vocabulary development. Table 1 includes only three examples of this. One is Joe (1998), where learners were asked to retell the content of a reading text, but there was neither a reading-only condition nor a vocabulary-focused learning condition to compare the attested vocabulary uptake to. Another example is Sun (2017), where students worked on tasks in which they shared their impressions and interpretations of picture-supported stories. Even though it was not stipulated if there were particular target words they should recycle from the stories, these tasks nonetheless generated vocabulary learning that was on par with what was achieved in a condition where the teacher invested time explaining all the unfamiliar words in the stories. The third example is Yang, Shintani, Li and Zhang (2017), which casts some doubt on the usefulness for vocabulary uptake of text-based output tasks if the vocabulary focus is not stipulated to learners. Learners in one of the conditions were asked to write a short essay (max 100 words) about the subject of a glossed text they had just read, but there was no real incentive for the learners to recycle content (and related vocabulary) from the input text, and it was found the participants did not incorporate any of the target words in their short essays. Unsurprisingly, then, the post-test scores turned out comparatively poor in this condition. However, a task where learners are requested to sum up or retell a text (e.g., Joe, 1998; Rassaei, 2017) naturally entails a greater need to include content matter of the original text and thus likely also the relevant vocabulary associated with it.

To the best of our knowledge, no close equivalents to the studies listed in Table 1 are as yet available for the case of vocabulary uptake from audio or audio-visual texts. Of potential relevance, however, are the following explorations of 'listen-and-do' activities, because they also illustrate the usefulness of output tasks. In Ellis and He (1999), learners followed aural directions

to put pieces of furniture (the referents of the target words) in their corresponding places in a picture. In two of the conditions, it was only the teacher who gave the directions. The teacher either explained the meaning of the target words from the start or waited for the students to ask for clarification. In a third condition, the target words were also first clarified, but then the students worked in pairs and took turns to not only follow but also give directions. This third condition—where the students themselves also used the target words productively—was found to bring about the best vocabulary gain. In an approximate replication, de la Fuente (2002) also had learners in two conditions only follow directions, while in a third condition the learners themselves became direction providers. Again, post-tests revealed that more target words were retained in this third condition. It is worth noting that listen-and-do tasks such as those used in Ellis and He (1999) and de la Fuente (2002) were carefully crafted so as to make the target words task-essential (i.e., without using/understanding them it would be almost impossible to fulfill the task; Loschky & Bley-Vroman, 1993). It is not immediately obvious how transferable such 'listen-and-do' designs are to the use of authentic audio or audio-visual materials.

Motivations for resorting to the TED Talks genre

TED Talks (https://www.ted.com) are freely available web-based conference presentations on a wide variety of professional and academic topics, delivered by experts who attempt to make these topics accessible for a general audience. They are relatively short (typically between four and twenty minutes) and thus generally manageable for classroom use in terms of length. It has been suggested (Coxhead & Walls, 2012) that one needs knowledge of the 4,000 most frequent word families in English to understand 95% of the running words of the average TED Talk—95% being the 'coverage' considered necessary for adequate listening comprehension (van Zeeland & Schmitt, 2012). While this may at first glance seem a rather high threshold, the deliberate efforts on the part of TED Talks presenters to clarify terms and elucidate concepts may render these videos comprehensible also for language learners who have not yet reached the level that provides 95% coverage. The presenters often use analogies, examples, visuals and props to clarify the concepts they introduce (Scotto di Carlo, 2014). One might argue that such clarifications will concern words of relatively low frequency and might therefore not be

of great importance for second language learners. Still, Laufer and Ravenhorst-Kalovski (2010), for example, have shown—in the context of academic reading—that learners' knowledge of low-frequency words can have a profound impact on text comprehension.

Given the general appeal of TED Talks it is not surprising that a growing number of language institutions worldwide—including the institution where we conducted our experiment—have started using TED Talks as a way of fostering learners' (academic) listening skills (e.g., Takaesu, 2013).

Why inserting an output activity may matter for incidental vocabulary acquisition

To reiterate, the study reported here examines the effect of asking learners to provide an oral summary (in their L2, i.e., English) of a TED Talk after they have watched it once and then to watch it again. The prediction that this summary activity will stimulate incidental vocabulary acquisition is in accordance with at least two tenets of Swain's Output Hypothesis (e.g., Swain, 1995). First, the output task is expected to promote noticing. When learners try to summarize the video content, they may experience lacunae in their lexical resources and this may subsequently prompt them to attend to the relevant lexis as they watch the video again. It is thus important for learners to revisit the input material (i.e., watch the TED Talk again, after the output task) to give them the opportunity to fill some of the gaps they noticed in their own resources. Second, in keeping with what Swain refers to as the hypothesis-testing affordance of an output task, the summary task (perhaps especially when the students know it is not a graded task) invites learners to experiment with using newly met words, and this may help to entrench these in memory. Again, re-visiting the input text can be useful for the learners to compare their own use of new words with how they are used in the 'model'. Put differently, while learners are likely to make guesses about the meaning of new words as they watch and listen to a TED Talk, the incentive for meaning guessing and for evaluating the plausibility of the guesses is probably stronger in a condition where learners seek to incorporate the new words in their own output.

The expectation that the summary task will positively influence vocabulary uptake from the TED Talk is also compatible with Laufer and Hulstijn's (2001) oft-cited construct of task-induced involvement—also known as the Involvement Load Hypothesis (Hulstijn & Laufer,

2001). According to this hypothesis, there are three components of engagement that help to predict vocabulary uptake from texts, i.e., need, search, and evaluation, and all three components appear to be present in the learning condition examined here. First, the summary task may create a need for the learners to mine the input text for lexis they deem useful for the performance of that task. Second, if learners experience lexical deficiencies during performance of the output task, this may prompt them to search for the missing pieces when the model text is revisited. Third, they are likely to compare their own attempted uses of the newly encountered words with how they are used in the model text, and this entails evaluation.

The use of newly learned words in output tasks is also included as a desirable feature of vocabulary learning activities in Nation and Webb's (2011) guidelines for 'Technique Feature Analysis', for two reasons. One is that it requires retrieval from memory—at least if the input text is not consulted during the output task, so words can simply be copied. The other is that learners' attempts at using words in new contexts—i.e., the 'generative' use of newly learned words (Joe, 1998)—can be particularly valuable for retention. It will therefore be worth examining in the below experiment not only to what extent students attempt to recycle novel words from the TED Talk in their own renderings of the content of the TED Talk, but also whether the sentential contexts in which the words are embedded differ from the original input text.

Research Questions

The primary research question addressed in this study is whether EFL students who sum up the content of a TED Talk and then watch it a second time pick up the meaning of more words from this input than same-profile students who watch the TED Talk twice without this output task. Secondary research questions that need to be asked, because they can help to account for an effect (if any) of the summary task, are whether the students (try to) use new words from the TED Talk in their oral summaries, and, if so, whether this (attempted) use is positively associated with the students' retention of the meaning of these words. As the students were told beforehand about the summary task, an additional question is whether the announcement of this particular output task enhances the students' processing of the content of

the TED Talk. This is relevant, because learners' successful vocabulary uptake from texts is known to hinge on their engagement with and comprehension of the texts (e.g., Krashen, 1996; Vidal, 2011).

Method

Participants

Two intact classes of Vietnamese learners (N = 64) of English as a foreign language participated in this quasi experiment. They were all 19 or 20 years of age and enrolled in a two-year intensive English language training program at a university in Viet Nam. The objective of the program was to develop the students' English proficiency to CEFR C1 level or IELTS overall band score of 6.5. According to an in-house English language proficiency test using CEFR as the underlying test construct, the students' proficiency in English at the time of the experiment was equivalent to CEFR B2+ or to IELTS overall band score of 5.5 (i.e., upper intermediate). In addition to the aforementioned proficiency test, we administered Nation and Beglar's (2007) Vocabulary Size Test (VST), which indicated that these learners had receptive knowledge of at least the 4,000 most frequent word families in English. As all the data in these tests were normally distributed, *t*-tests for independent samples were run to compare the two groups' L2 proficiency test scores and their VST scores. No difference was found in either their L2 proficiency scores or their VST scores, with $t = 0.33$ ($p = .74$) and $t = 0.34$ ($p = .73$), respectively.

Input material and target words

Because the existing EFL program already included a substantial reading component but lacked an equally well developed listening component, we opted to use a TED Talk without captions and without giving the students access to the transcript. Research suggests that learners allocate a lot of their attention to captions (Winke, Gass, & Sydorenko, 2013) and tend to rely on their better developed reading skills in L2 to bypass the need to process the listening input (Garza, 1991). We recognize that several studies have demonstrated that captions assist both text comprehension and vocabulary uptake (see Montero Perez, Van Den Noortgate, & Desmet,

2013, for a meta-analysis). Nevertheless, as one of the objectives of the language program was to equip students with the skills to cope with authentic listening input, unassisted by written text, using a TED Talk without captions was deemed pedagogically and ecologically justified. This does not mean that written language was totally absent from the input, though, because some of the visuals used in the selected TED Talk included written words—something we will return to later.

Sixteen learners who had roughly the same level of L2 proficiency and vocabulary knowledge as those recruited for the actual study were invited to rate a selection of TED Talks on various topics with regard to their comprehensibility and general appeal. Based on these learners' feedback, a 12-minute talk about marine bioluminescence was chosen. According to Nation and Heatley's RANGE program (2002), 20 words in this talk belonged to frequency bands beyond the most frequent 4,000 word families in English, and these were selected as the potential target items. Choosing these words minimized the likelihood that the participants might already know these items prior to the experiment. It is worth mentioning that these 20 words were included in the tests as the *potential* target words in this study. The actual target words for comparing participants' learning gains from the pre-test to the post-tests were selected based on the participants' pre-test performance on these items (see below).

Vocabulary knowledge measure

The vocabulary knowledge measure in this study was a word-meaning recall test, i.e., a test of receptive word knowledge where the English target word is presented as prompt and the learner is asked to supply its meaning. The word-meaning recall format has been successfully used in the studies by van Zeeland and Schmitt (2013) and Brown et al. (2008), for instance. However, the tests in these studies share a potential limitation in that all test prompts were presented in written form even though the participants were not exposed to this in the input materials. Sydorenko (2010) and Winke et al., (2010), have argued that the vocabulary test performance may be susceptible to the mode of presentation of the test prompts if this is not congruent with that of the input texts. Therefore, our test used the spoken rather than written word forms as the test prompts, because it could not be taken for granted that learners would

always recognize the words they had been exposed to if these were presented to them in written form.

To prepare the test, a male native speaker of English was invited to read aloud and audio-record the 20 potential target words with an interval of 20 seconds between the words. The result of a pilot study showed that this interval was long enough for L2 learners to write down their responses. The participants were invited to provide the meaning of the target words in Vietnamese (L1), prompted by the audio-recorded English word. This test was given to the participants three times: two weeks before the experiment, at the end of the experiment, and two weeks after the experiment (henceforth referred to as the pre-test, the immediate post-test, and the delayed post-test, respectively). Each time the test was administered, the order of the test items was randomized to avoid item sequence effects.

In the scoring procedure, two different rating scales were used. The first rating scale allowed measurement of partial vocabulary knowledge. On this scale, 0, 0.5, and 1 point was awarded to any incorrect, partially correct, and fully correct response, respectively.
For example, such responses as "*sinh vật biển*" (sea animals), "*phát quang*" (luminescence), and "*phát quang sinh học*" (bioluminescence) to the item "*bioluminescence*" were awarded 0, 0.5, and 1 point, respectively. The second rating scale was stricter as it only awarded points to fully correct responses. Therefore, only the last response "*phát quang sinh học*" (bioluminescence) in the example above was given credit. The first author (a Vietnamese-English bilingual) first scored all the responses. To establish reliability, two other experienced Vietnamese teachers of English independently scored 10% of all test papers as well. No differences were found between the three assessors' scores.

Text comprehension measure

According to Lynch (2009), conference presentations such as TED Talks belong to the realm of one-way academic listening, as they bear resemblance to university lectures. To gauge learners' comprehension of this kind of material, guidelines proposed in the literature on L2 academic listening are therefore useful. Wagner (2002) describes the construct of L2 academic listening with reference to the well-established two-pronged model of top-down and bottom-up

processing. After revisiting the tests used in earlier work on academic listening, he argues that the most valid way of assessing L2 academic listening comprehension is a combination of tasks that require learners to: (a) identify the purpose and main ideas of the text, (b) make inferences about implicit meaning, logical links between ideas, and speaker attitude, and (c) recognize important supporting ideas and factual information. Thus, we designed a text comprehension test with these components in mind.

Fifteen content questions in a YES/NO format were used to gauge the extent to which the participants understood and remembered the video content. Three experienced EFL/ESL teachers (one a native speaker and the others non-native speakers of English) were invited to watch the selected video to determine the aforementioned three types of information. Fifteen sets of information proposed by all three teachers were then chosen to design the test. These questions were all presented in the form of statements for the learners to decide whether or not they reflected the video content, by writing "Yes" or "No" next to each statement. Below are three examples of these statements: "*One reason why sea animals produce light is to attack their enemies"* (main idea), "*There is evidence that living things on other planets can also produce light*" (implicit), and "*A Humboldt squid has even tried to rescue this electronic jellyfish*" (factual). The internal consistency of the test was $\alpha = .92$. It should also be noted that no target words were used in these statements.

Procedure

The participants in both treatment conditions were asked to watch the TED Talk twice with a view to answering content questions after the second viewing. They were encouraged to take notes.

In one treatment condition—henceforth the 'summary' condition—they were told they would also be required to orally sum up (in English) the content of the TED Talk in five minutes after the first viewing. These oral summaries were recorded (for transcription and screening for attempted uses of target words). The comparison group were told they would be given five minutes after the first viewing to revise and organize their notes so as to make it easier for them to verify and add information during the second viewing. 'Time on task' was thus the same in

both conditions. After the second viewing all the participants were given the 15 Yes/No content questions. They were allowed to consult their notes to answer these questions (but they did not have access to the TED Talk at this stage).

The vocabulary post-test was administered immediately after this and again after a two-week interval. The students were not forewarned of the vocabulary tests.

Analysis

The purpose of the vocabulary pre-test in this study was to establish a set of target words which were unknown to the participants prior to the experiment. Following the example of studies described in Ellis (1999), two words that were already known by more than 10% of the participants were removed from the final analyses—i.e., although they were included in the tests, they were no longer considered 'target words'. The number of the target words was thus reduced from 20 to 18. A list of these 18 target words, their number of in-text occurrences, and a link to the video of the TED Talk are provided in Appendix.

In what follows the terms uptake and retention will be used to refer to the learning outcome as evidenced in the immediate post-test and the delayed post-test, respectively. Uptake is calculated by subtracting the pre-test scores from the immediate post-test scores while retention is determined by subtracting the pre-test scores from the delayed post-test scores.

As all test data were normally distributed, we used the conventional ANOVA tests followed by Tukey HSD calculations, Cohen's effect size *d* and *t*-tests to compare the size of vocabulary uptake and retention within and between the conditions. In addition, Pearson correlation was computed to examine the association between vocabulary gain and text comprehension. A two-tailed *p* value of .05 was set as a threshold for significance in all the tests.

Results

The primary research question addressed in this study is whether the insertion of an oral output task (more precisely, a summary task) between two viewings of a TED Talk positively influences vocabulary acquisition, without any intervention to direct learners' attention to target

words or to clarify word meanings (beyond clarifications given in the TED Talk itself) or to stipulate they should use particular words in their summaries.

Table 2 presents the descriptive statistics for the scores obtained under the two treatment conditions on the partial vocabulary knowledge rating scale. Recall that, according to this scoring, responses were given 0, 0.5 or 1 point. Given that 18 of the original set of 20 words were retained as actual target words, the maximum obtainable score is 18.

Table 2: Mean gains including credit for partial knowledge (SDs in parentheses)

Groups	Pre-test scores	Immediate post-test scores	Delayed post-test scores	Uptake scores	Retention scores
Comparison	0.19 (0.40)	4.61 (2.11)	3.50 (1.68)	4.42 (1.92)	3.31 (1.55)
Summary	0.22 (0.49)	8.16 (4.71)	6.72 (3.78)	7.94 (4.42)	6.50 (3.50)

To examine whether the participants' knowledge related to the meaning of the target words differed between conditions prior to the experiment, a *t*-test for independent samples was carried out. No difference was found: $t = 0.28$ ($p = .78$). It is therefore safe to attribute any difference in the vocabulary gain between the groups to the learning conditions. One-way ANOVA tests for correlated samples were carried out to gauge the difference in the learners' scores across the three testing times – the pre-test, the immediate post-test, and the delayed post-test. A significant difference was found in both conditions, where the post-test scores were significantly different from the pre-test scores. In the summary condition, there was no significant difference between the participants' immediate and delayed post-test scores ($p = .23$), suggesting that knowledge was well retained over time. This does not hold true to the same extent in the comparison condition, where the difference between the immediate and delayed post-test scores does reach significance ($p < .05$). *T*-tests for independent samples were run to compare the size of vocabulary uptake and retention between the two conditions. This showed that both the uptake and the retention were significantly larger in the summary condition than that in the comparison condition, with $t = 4.13$ ($p < .001$) and $t = 4.71$ ($p < .0001$), respectively.

Cohen's effect size *d* was also computed to examine the difference in lexical uptake and retention between the two conditions. Such a difference was consistently found to be large with $d = 1.03$ in the case of the uptake scores and $d = 1.18$ in the case of the retention scores.

A similar picture emerges from the participants' test scores as per the strict (and thus dichotomous) scoring, as shown in Table 3.

Table 3: Mean gains of precise knowledge (SDs in parentheses)

Groups	Pre-test scores	Immediate post-test scores	Delayed post-test scores	Uptake scores	Retention scores
Comparison	0.19 (0.40)	1.84 (0.85)	1.75 (0.84)	1.65 (0.70)	1.56 (0.67)
Summary	0.22 (0.49)	2.72 (1.57)	2.69 (1.51)	2.50 (1.32)	2.47 (1.27)

No difference was found between the two groups' pre-test scores also according to the dichotomous scoring of responses: $t = 0.28$ ($p = .78$). To examine whether the participants' scores significantly improved from the pre-test to the post-tests, a one-way ANOVA test for correlated samples was run for each condition. Both groups significantly enhanced their scores from the pre-test to the post-tests with $F(2, 93) = 52.47$ ($p < .000$), and $F(2, 93) = 39.55$ ($p < .000$). There was no significant difference between the immediate and delayed post-test scores in either condition. *T*-tests for independent samples were also run to compare the vocabulary uptake and retention between the conditions. The vocabulary uptake and retention in the summary condition were again significantly greater than those in the comparison condition, with $t = 3.19$ ($p < .01$) and $t = 3.47$ ($p < .01$), respectively. To investigate the difference in the vocabulary gain between the two treatment conditions, Cohen's effect size *d* was also calculated. This difference was large in the case of lexical uptake with $d = 0.81$. The same holds true for the case of lexical retention with $d = 0.90$.

If the summary condition brought about better vocabulary uptake and retention, the question needs to be asked whether this may be accounted for by students' attempted use of target words (and subsequent verification against the 'model' input). To explore this possibility,

the recorded summaries were screened for the use of target words. Altogether we found five out of the 18 target words in the collection of summaries: *lure* (16 tokens), *bioluminescence* (14 tokens), *plunge* (6 tokens), *bait* (5 tokens), and *breath-taking* (2 tokens). Eighteen of the 32 learners used at least one target word in their summaries. These words were unfamiliar to these learners prior to the treatment, according to their pre-test data. Nevertheless, they were used in their summaries in grammatically and semantically appropriate ways, and this in utterances which were no verbatim duplications of utterances used by the TED Talk presenter. Below is one example of how target words were used in the summaries:

S20: "*In two examples, she used an optical **lure**. When this optical lure started producing light, four or five fish came along and they started to produce light and together lighting up for communication. And then a shark came. She said she didn't know it was because of this light or the **bait** that she put it in the optical **lure** that **lured** the shark.*"

Of direct relevance for the question of whether the summary task benefited vocabulary learning is the finding that, when participants used a target words in their summaries, they also successfully recalled it in the two post-tests.

In the summaries, we also found 42 instances of conspicuously long (more than two seconds) pauses where a target word would have been appropriate. All these pauses occurred in the middle of a clause, a hesitation which is a likely sign that learners were looking for a word (Tavakoli, 2011), but then failed to retrieve it from memory. After the pause, the participant would either resort to a substitution or a paraphrase, as in "*They set [a long pause] a machine; the machine can produce light and therefore it attracts many fish like an octopus and shark*", presumably when this learner was trying to remember "optical *lure*". Interestingly, 22 of these 42 likely manifestations of unsuccessful retrieval attempts were followed by successful word recall in the post-tests. It is therefore plausible that the failures to recall words from the TED Talk prompted the learners to take notice of these elusive target words as they engaged with the TED Talk a second time—consistent with the 'noticing' tenet of the Output Hypothesis.

A complementary account for the positive influence of the output task on vocabulary acquisition is its association with text comprehension. On the listening comprehension test, the participants in the summary and the comparison group obtained a mean score of 9.31 (SD = 1.91)

and 7.25 (*SD* = 2.24) out of the maximum score of 15 points, respectively. A *t*-test yields a significant difference: $t = 3.54$ ($p < .001$). Pearson correlation coefficients were computed to examine the association between the text comprehension scores and the vocabulary gain scores. On the partial knowledge rating scale, the participants' text comprehension scores were positively correlated with their vocabulary uptake and retention scores with a coefficient of *r* .33 and .36 ($p < .01$), respectively. These coefficients in the case of the precise knowledge rating scale were, also in that order, *r* .28 and .31 ($p < .05$).

Discussion and Conclusion

Regarding the primary research question, when learners were asked to sum up the content of the TED Talk before watching it a second time, they gained knowledge of three (16.67%) and eight (44.44%) out of the 18 target words on average according to the strict scoring method and the scoring method that gave credit for partial knowledge. While arguably modest, this vocabulary gain nonetheless surpassed the gain attested in the comparison condition, where learners also watched the TED Talk twice but without insertion of an oral output task. In the latter condition, the students learned on average two (11.11%) and four (22.22%) words according to the two scoring protocols. The students in the summary condition also obtained significantly better scores on the text comprehension test than the comparison group. It is conceivable that the announced summary task helped the learners process the input material with a clearer purpose in mind than only the announcement that content questions would follow. It is also possible that the actual retelling helped them to consolidate the information gathered from the first viewing (Wittrock, 1991). In addition, if the learners felt their summary was incomplete or perhaps not fully accurate, this may have prompted them to seek the missing information in the video during the second viewing.

As far as the secondary research questions are concerned, the summary task certainly created a near-immediate opportunity for the learners to try using newly met words, and—unlike what was attested in Yang, et al. (2017), where the chosen output task was more loosely connected to the input text—the students made use of this opportunity. Crucially, when learners used target words in their summaries, they also managed to recall the meaning of these words at

least up to two weeks later. This may attest to the usefulness of retrieval and of the generative use of new words, i.e. two desirable features of vocabulary learning activities according to Nation and Webb's (2011) Technique Feature Analysis. This finding is also in keeping with Laufer and Hulstijn's (2001) Involvement Hypothesis, because the active use of a newly met word entails a certain evaluation on the part of the learner of its accuracy and appropriateness in the given context. When the word is subsequently re-encountered in the 'model' text, further evaluation of one's own use of the word is likely to happen.

Moreover, even where learners tried but failed to incorporate target words in their oral summaries, the output task seems to have been indirectly conducive to learning. That, at least, is suggested by the many episodes of hesitations suggesting efforts of retrieval that were followed by successful recall in the post-tests. The latter finding is in accordance with the Output Hypothesis and again also illustrates the usefulness of revisiting the input text which a given output task is based on so as give learners the opportunity to mine that input for elements they felt lacking in their expressive resources.

It is worth emphasizing that the study reported here was a classroom-based experiment. It is unlikely that learners will often engage in content-retelling after they have watched a video *outside* the classroom—unless they really feel the desire to tell someone who has not watched it. As a classroom activity (or as an activity incorporated in an online interactive course), it is relatively easy to implement, however. While (for the sake of experimental control) the participants in the present study all watched the same video and recorded their summaries individually in the form of monologues, in more genuinely communicative practice students could be asked to watch different videos on a related topic and then be paired up to exchange information about them.

Although the chosen output task enhanced vocabulary learning from the TED Talk, it needs to be acknowledged that the outcome in terms of the number of new words actually learnt still remains modest. Many factors may help account for this. First, the study material in this experiment was authentic in the sense that it was produced without L2 learners in mind. Such characteristics of authentic materials as natural and fast speech must be challenging to learners. It is worth noting in this regard that much previous research on incidental vocabulary acquisition

used either special-purpose instructional materials, like graded readers (e.g., Brown et al., 2008; Horst et al., 1998) or materials developed specifically for research purposes (e.g., van Zeeland & Schmitt, 2013; Vidal, 2011). For instance, as the video in this experiment was not modified, word repetition in the input text was not manipulated. Consequently, all but one of the 18 target words ("*bioluminescence*") occurred only once or twice in the text, which is much lower than the 10+ or 20+ times suggested by van Zeeland and Schmitt (2013, p. 621) or Brown et al. (2008, p. 153) as necessary for incidental vocabulary acquisition to occur. Second, although learners with a similar profile as the actual participants had suggested that the selected TED Talks video was suitable for use in this study (see section Input Material above), the comprehension test scores indicated that its content was far from straightforward. At the same time, the TED Talk chosen for the experiment illustrates some of the affordances of the genre for L2 vocabulary acquisition more generally. For example, one of the words that was relatively well remembered was "*bioluminescence*". In the TED Talk, the meaning of this word was clearly denoted via a verbal explanation "*So it is a little appreciated fact that most of the animals in our ocean make light. I have spent most of my career studying this phenomenon called bioluminescence*" and a series of corresponding footages. This item might also have stood a better chance of learning as it was essential for text comprehension and occurred six times in the given passage. Another example was the word "*lure*" in the phrase "*optical lure*". In the video, the written form of this phrase was put on the screen and its meaning was both visually and verbally demonstrated as in "*This is an optical lure that I have used. We call it the electronic jellyfish. It is just 16 blue LEDs that we can program to do different types of [light] display*."

It would be interesting to be able to compare the learning rate observed in the present study to the learning rates reported in studies such as those listed in Table 1 where the participants engaged in more deliberate post-reading work on vocabulary. While such comparisons should in any case be interpreted very cautiously because of differences in materials, procedures, participants and so forth, they are particularly difficult here owing to the diversity of test formats used. One study in Table 1, Hulstijn and Laufer (2001), used a meaning-recall test like we did and also gave credit for partially correct responses. The task that led to the best learning gains in two trials of the same experiment (conducted with different student

populations) was writing a text incorporating ten target words from a glossed reading text. Total time spent on the reading and writing tasks was 70-80 minutes. The delayed post-test scores were 2.6 in one trial and 3.7 in the other. The success rate attested in the experiment we have reported here does not compare unfavorably to this considering that it was obtained through classroom procedures that took only about 40 minutes. Also Lu (2013) included a meaning-recall test (but responses were scored only dichotomously). The students read a story containing nine glossed target words, and then performed tasks with a focus on these nine words, such as summary writing incorporating the words (which took close to 40 minutes in addition to the time spent on reading the story) and completing three gapped sentences for each word (which took close to 20 minutes in addition to the time spent on the story). On average, the students in these two conditions recalled the meaning of two of the nine target words in the delayed post-test. Again, the learning we have observed in our experiment, without input modification and without instruction to engage students in deliberate word learning, seems not to compare unfavorably to this.

As with other studies of vocabulary uptake from textual input, it needs to be borne in mind that the gains observed in a vocabulary test on a selected set of target words is likely to underestimate the amount of acquisition that actually took place. After all, the participants may have acquired, fine-tuned, or strengthened knowledge of other words than the 18 target words we selected. Besides, there are of course multiple other ways beyond word learning in which the participants are likely to have benefited from engaging with the TED Talks video and from the meaning-focused oral output activity.

Even as regards the 18 target words we used to estimate vocabulary learning, it needs to be borne in mind that these were selected simply because they were unknown to the participants. Whether the students considered many of them to be of high utility and thus worthy of remembering is far from certain—but this is no different from numerous previous studies where target words were selected by virtue of their low frequency. It also needs to be acknowledged that we used only one test format to measure knowledge gains, and this was a rather challenging format. For example, if, apart from the meaning recall format, we had administered a less

challenging meaning recognition test (in a multiple-choice format), this could have yielded evidence of learning which the recall test may have failed to reveal.

The current study focused on vocabulary uptake through an input-output-input cycle, but using this cycle is likely to be beneficial in additional ways. For example, it can also help learners to develop their L2 listening metacognition strategies. Specifically, when they are required to summarize listening content, learners need to synthesize information from their notes. This synthesis helps them to notice gaps in their previous interpretation of input content, which, in turn, prompts them to direct their attention to relevant information as they listen to the same input a second time. Put differently, the input-output-input cycle provides them with an opportunity to monitor their listening process, evaluate the listening outcome, and search for suitable solutions to the listening problems that may have experienced (Vandergrift & Goh, 2012).

There are inevitably several limitations to the study we have reported here. One limitation is that only one type of audio-visual input and only one type of output task were put to the test. It is worth examining whether similar trends might be observed (and similar benefits reaped) in conceptual replications with other video materials (including captioned ones) and with the implementation of other text-based output tasks, especially ones that stimulate learner-learner interaction instead of the unidirectional summary task tried here. A likely factor in the effectiveness of text-based output tasks for vocabulary uptake (when it is not made mandatory to use particular words) is the extent to which the learners will find words from the text useful or even essential to complete the output task they have been set. Even in the case of a summary or retell task, it is clear that non-key concepts (and associated words) may be avoided. In our data such words as "*barbel*", "*mesh*" and "*morsel*" were not recycled from the text and consequently stood a poor chance of being recalled later. Besides, as we have seen, even in the case of key concepts, learners may be able to resort to strategies such as substitution and paraphrase. Another limitation to this study is that it did not include another treatment condition where an output activity followed the initial viewing but without a subsequent viewing. This could have helped to determine to what extent the (announced) output task as such promotes vocabulary uptake, even in the absence of a second viewing.

References

Baltova, I. (1994). The impact of video on the comprehension skills of core French students. *Canadian Modern Language Review, 50*, 507–31.

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

Chen, C., & Truscott, J. (2010). The effects of repetition and L1 lexicalization on incidental vocabulary acquisition. *Applied Linguistics, 31*, 693–713.

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

Coxhead, A. J., & Walls, R. (2012). TED Talks, vocabulary, and listening for EAP. *The TESOLANDZ Journal, 20*, 55-58.

de la Fuente, M. J. (2002). Negotiation and oral acquisition of L2 vocabulary. *Studies in Second Language Acquisition*, *24*(1), 81-112. doi: 10.1017/S0272263102001043.

Eckerth, J., & Tavakoli, P. (2012). The effects of word exposure frequency and elaboration of word processing on incidental L2 vocabulary acquisition through reading. *Language Teaching Research, 16*, 227–252. doi: 10.1177/1362168811431377.

Ellis, R. (1999). *Learning a second language through interaction*. Amsterdam: John Benjamin Publishing Co.

Ellis, R., & He, X. (1999). The roles of modified input and output in the incidental acquisition of word meanings. *Studies in Second Language Acquisition, 21*(2), 285-301. doi: 10.1017/S0272263199002077.

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.

Hu, H-C. M., & Nassaji, H. (2016). Effective vocabulary learning tasks: Involvement Load Hypothesis versus Technique Feature Analysis. *System, 56*, 28–39.

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

Hulstijn, J. H., & Laufer, B. (2001). Some empirical evidence for the involvement load hypothesis in vocabulary acquisition. *Language Learning, 51,* 539–558.

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

Joe, A. (1998). What effects do text-based tasks promoting generation have on incidental vocabulary acquisition? *Applied Linguistics, 19,* 357–377.

Jung, J. (2016). Effects of glosses on learning of L2 grammar and vocabulary. *Language Teaching Research,* 20, 92 –112. doi: 10.1177/1362168815571151.

Keating, G. D. (2008). Task effectiveness and word learning in a second language: The involvement load hypothesis on trial. *Language Teaching Research, 12,* 365–386. doi: 10.1177/1362168808089922.

Kim, Y. (2008). The role of task-induced involvement and learner proficiency in L2 vocabulary acquisition. *Language Learning, 58,* 285–325.

King, J. (2002). Using DVD feature films in the EFL classroom. *Computer Assisted Language Learning, 15,* 509–523.

Ko, M. H. (2012). Glossing and second language vocabulary learning. *TESOL Quarterly, 46,* 56−79.

Kost, C. R., Fost, P., & Lenzini, J. J. (1999). Textual and pictorial glosses: Effectiveness of incidental vocabulary growth when reading in a foreign language. *Foreign Language Annals, 32,* 89–113.

Krashen, S. D. (1996). The case for narrow listening. *System, 24,* 97-100. doi: 10.1016/0346-251X(95)00054-N.

Laufer, B. (2003). Vocabulary acquisition in a second language: Do learners really acquire most vocabulary by reading? Some empirical evidence. *The Canadian Modern Language Review, 59,* 567–587.

Laufer, B., & Hulstijn, J. (2001). Incidental vocabulary acquisition in a second language: The construct of task-induced involvement. *Applied Linguistics, 22,* 1–26. doi: 10.1093/applin/22.1.1.

Laufer, B., & Ravenhorst-Kalovski, G. C. (2010). Lexical threshold revisited: Lexical text coverage, learners' vocabulary size and reading comprehension. *Reading in a Foreign Language, 22,* 15–30.

Laufer, B., & Rozovski-Roitblat, B. (2011). Incidental vocabulary acquisition: The effects of task type, word occurrence and their combination. *Language Teaching Research, 15,* 391–411.

Laufer, B., & Rozovski-Roitblat, B. (2015). Retention of new words: Quantity of encounters, quality of task, and degree of knowledge. *Language Teaching Research, 19,* 687–711.

Lee, S. & Pulido, D. (2017). The impact of topic interest, L2 proficiency, and gender on EFL incidental vocabulary acquisition through reading. *Language Teaching Research, 21*. doi: 10.1177/1362168816637381.

Loschky, L., & Bley-Vroman, R. (1993). Grammar and task-based methodology. In G. Crookes & S. M. Gass (eds.), *Task and language learning: Integrating theory and practice* (pp. 123–167). Clevedon: Multilingual Matters.

Lu, M. (2013). Effects of four vocabulary exercises on facilitating learning vocabulary meaning, form, and use. *TESOL Quarterly, 47,* 167–176.

Lynch, T. (2009). *Teaching second language listening*. Oxford: Oxford University Press.

Min, H.T. (2008). EFL vocabulary acquisition and retention: Reading plus vocabulary enhancement activities and narrow reading. *Language Learning, 58,* 73–115.

Montero Perez, M., Van Den Noortgate, W., & Desmet, P. (2013). Captioned video for L2 listening and vocabulary learning: A meta-analysis. *System, 41,* 720–739. doi: 10.1016/j.system.2013.07.013.

Nation, I. S. P. (2007). The four strands. *Innovation in Language Learning and Teaching, 1,* 1–12.

Nation, I. S. P., & Beglar D. (2007). A vocabulary size test. *The Language Teacher, 3,* 9–13.

Nation, I. S. P., & Heatley A. (2002). *Range: A program for the analysis of vocabulary in texts*. http://www.victoria.ac.nz/lals/about/staff/paul-nation

Nation, I.S., & Webb, S.A. (2011). *Researching and analyzing vocabulary*. Boston, MA: Heinel.

Neuman B., & Koshinen, P. (1992). Captioned television as comprehensible input: Effects of incidental word learning from context for language minority students. *Reading Research Quarterly, 27,* 96–106.

Paribakht, S., & Wesche, M. (1997). Vocabulary enhancement activities and reading for meaning in second language vocabulary acquisition. In: J. Coady, & T. Huckin (Eds.), *Second language vocabulary acquisition: A rationale for pedagogy* (pp.174–200). New York: Cambridge University Press.

Peters, E. (2012). The differential effects of two vocabulary instruction methods on EFL word learning: A study into task effectiveness. *International Review of Applied Linguistics, 50,* 213–238. doi: 0019042X/2012/050-213.

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*, 113–151.

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

Pulido, D. (2007). The relationship between text comprehension and second language incidental vocabulary acquisition: A matter of topic familiarity? *Language Learning, 57*, 155–199.

Qian, D. D. (2002). Investigating the relationship between vocabulary knowledge and academic reading performance: An assessment perspective. *Language Learning, 52,* 513–536.

Rassaei, E. (2015). Effects of three forms of reading-based output activity on L2 vocabulary learning. *Language Teaching Research, 21*, 76 –95. doi: 10.1177/1362168815606160

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

Rodgers, M. P. H., & Webb, S. (2011). Narrow viewing: The vocabulary in related television programs. *TESOL Quarterly*, *45*(4), 689–717. doi: 10.5054/tq.2011.268062.

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.

Scotto di Carlo, G. (2014). The role of proximity in online popularizations: The case of TED talks. *Discourse Studies, 16,* 591–606.

Stæhr, L. S. (2008). Vocabulary knowledge and advanced listening comprehension in English as a foreign language. *Studies in Second Language Acquisition, 31*, 577–607.doi:10.1017/S0272263109990039.

Sun, C-H. (2017). The value of picture-book reading-based collaborative output activities for vocabulary retention. *Language Teaching Research, 21*, 96 –117. doi: 10.1177/1362168816655364.

Swain, M. (1995). Three functions of output in second language learning. In G. Cook & G. Seidhofer (Eds.), *Principles and practices in applied linguistics: Studies in honour of H. G. Widdowson,* (pp. 125-144). Oxford: Oxford University Press.

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

Takaesu, A. (2013). TED Talks as an extensive listening resource for EAP students. *Language Education in Asia, 4*, 150-162.

Tavakoli, P. (2011) Pausing patterns: differences between L2 learners and native speakers. *ELT Journal*, 65(1), 71-79. doi: 10.1093/elt/ccq020.

Vandergrift, L., & Goh, C. C. (2012). *Teaching and learning second language listening: Metacognition in action*. New York, NY: Routledge.

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

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. (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.

Wagner, E. (2002). Video listening tests: A pilot study. *Working Papers in TESOL and Applied Linguistics, 2*, 1-39.

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*, 130–160.

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

Webb, S. (2008). The effects of context on incidental vocabulary learning. *Reading in a Foreign Language, 20*, 232–245.

Webb, S. (2014). Repetition in incidental vocabulary learning. In C. A. Chapelle (Ed.), *The Encyclopedia of Applied Linguistics* (pp. 1–6). New York: John Wiley and Sons.

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). London: Routledge.

Webb, S., & Chang, A. C-S. (2015). Second language vocabulary learning through extensive reading with audio support: How do frequency and distribution of occurrence affect learning? *Language Teaching Research, 19*, 667–686. doi: 10.1177/1362168814559800.

Winke, P., Gass, S., & Sydorenko, T. (2013). Factors influencing the use of captions by foreign language learners: An eye-tracking study. *The Modern Language Journal*, *97*(1), 254-275. doi: 10.1111/j.1540-4781.2013.01432.x.

Winke, P., Gass, S., & Sydorenko, T. (2010). The effects of captioning videos used for foreign language listening activities. *Language Learning & Technology, 14*, 65–86.

Wittrock, M. C. (1991). Generative Teaching of Comprehension. *Elementary School Journal*, *92*(2), 169-84. doi: 10.1086/461686.

Yang, Y., Shintani, N., Li, S., & Zhang, Y. (2017). The effectiveness of post-reading word-focused activities and their associations with working memory. *System, 70*, 38–49. DOI: 10.1016/j.system.2017.09.012.

Zou, D. (2017). Vocabulary acquisition through cloze exercises, sentence-writing and composition-writing: Extending the evaluation component of the involvement load hypothesis. *Language Teaching Research, 21*, 54–75. doi: 10.1177/1362168816652418

Appendix

Link to the selected video	Target words and their frequency of occurrence
https://www.ted.com/talks/edith_widder_the_weird_and_wonderful_world_of_bioluminescence	*propagate (1), plunge (2), mesh (2), bait (2), breath-taking (2), morsel (1), trawling (1), plankton (2), contemplative (1), plume (2), barbel (2), barrage (2), unobtrusive (1), entranced (1), lure (1), foolhardy (1), junky (1), bioluminescence (6)*