The effect of output and audiovisual input on L2 learners' oral productive vocabulary use

MA Dissertation
Author: Anna-Liisa Seveljova
Supervisor: Dr. Ana Fernández Montraveta
MA in Advanced English Studies
July 2024

Acknowledgments

First and foremost, I would like to express my sincere gratitude to my supervisor, Dr. Ana Fernández Montraveta, for her guidance, feedback and flexibility during the process of writing this MA dissertation. I am also grateful to my professors and colleagues at the UAB for inspiring me and leading me to this fascinating research topic. A special mention also goes to my dear friend Andy, who recorded the narration for the experiment, bringing the story to life.

Finally, I would like to thank teacher Sara Gómez, who kindly offered me the opportunity to conduct the experiment with her students, as well as all the learners who volunteered their time and effort. Without their participation, this MA thesis would not have been possible.

Table of Contents

Index of Tables

Abstract

In English Foreign Language (EFL) classes, it is crucial to help learners foster their productive use of lexis. It has been suggested that foreign language (L2) input could be used to this end, especially when it is paired with activities which may help learners notice what linguistic resources they are lacking. Thus, the current study analyzes learners' productive vocabulary use after being exposed to L2 audiovisual (AV) input to examine to what extent learners incorporate lexical items from input into their subsequent oral output. Additionally, the study explores whether adding an oral output phase before input exposure could facilitate noticing of necessary lexical items, and therefore their use. In order to answer these questions an empirical study was carried out with two groups of adult B1 level English learners. Both groups watched a narrated story and retold it, but one group also told the story with the help of images before input exposure. The proportion of shared words between the AV input and the oral tasks in each output phase indicates that the AV input significantly influenced the participants' lexical choices. Additionally, an added output phase seems to have a positive effect on noticing lexical items in the input. These findings suggest that using AV input may be an effective pedagogical tool to promote the productive use of lexis in the EFL classroom.

1   Introduction

Vocabulary knowledge is undoubtedly one of the pillars of English proficiency as it is fundamental to all language skills. Vocabulary size is not only strongly associated with the receptive skills of reading and listening, but it also plays a vital role in the two productive skills of speaking and writing (for an overview of studies see Qian and Lin, 2019). For foreign language (L2) learners, their lexical competence is a determining factor in achieving successful comprehension and communication and should therefore be one of the priorities of both language instructors and learners.

A similar distinction between receptive and productive ability has traditionally been made in vocabulary research, where receptive mastery refers to being able to understand lexical items when listening or reading, while productive mastery means being able to produce lexical items when speaking or writing (Schmitt, 2014). There is a consensus that learners have notably less productive than receptive vocabulary (Webb, 2008). As this may hinder effective communication, promoting learners' productive use of lexis in the EFL classroom becomes key. The current MA thesis aims to shed light on whether this could be done with the help of AV input by examining the impact of a narrated L2 story on learners' vocabulary use in their own retelling of the story.

This paper is organized as follows. First, the relevant literature is presented, with a focus on studies addressing the productive use of vocabulary, the role of input in L2 vocabulary use and development, noticing, and the role of output in noticing. Next, research questions and hypotheses are proposed. The methodology section follows, describing the participants, input materials, tasks, and procedure. The results of the quantitative analysis are then presented. The subsequent discussion section interprets the results, through both a quantitative and qualitative lens. Finally, the conclusion section highlights the most important findings and their

pedagogical implications, acknowledges limitations and discusses potential avenues for further research.

2   Literature review

2.1   L2 learners' productive vocabulary

While it has been established that learners generally know words receptively before being able to produce them (Laufer and Paribakht, 1998), it remains unclear exactly how and when this shift from receptive to productive ability happens. It has been proposed that as learners acquire more information about a word, their initial ability to recognize and comprehend it progressively evolves into productive mastery, where they can actively use it themselves (Melka, 1997, quoted in Schmitt, 2014). Schmitt (2019) claims that while learning a word from no knowledge to receptive mastery is relatively easy, the real challenge lies in enhancing the word knowledge enough to reach productive mastery.

It is not surprising that producing words in speech or writing is more challenging for learners than understanding them while reading or listening. During reading and listening comprehension activities, learners can manage by merely knowing the meaning of a particular word or may even use context to help them. In contrast, to be able to use it appropriately in their own output, they would also need to know the correct spelling, pronunciation, collocation, grammatical functions, among other aspects (Schmitt, 2019). Therefore, it is crucial to provide L2 learners with additional support to foster their productive vocabulary use.

It has been observed that learners may sometimes be able to provide a lexical item when prompted (i.e. this word is indeed a part of their productive vocabulary), yet they might never

use this item voluntarily. This led some researchers to propose a further distinction within productive ability. Laufer and Paribakht (1998) differentiate between *free* and *controlled active* vocabulary – the former refers to words that learners produce when they are free to choose which words to use (e.g. when telling a story), and the latter refers to specific words that learners can produce when prompted to do so (e.g. in a gap-fill exercise). The current study investigates free active vocabulary through an oral story retelling task, allowing participants to choose their words freely without the requirement to use particular lexical items. While the pictures provided to participants in the experiment try to cue them to use specific words, the choice is ultimately theirs, since they may omit some details or use an alternative lexical item.

When given the freedom to choose, learners often avoid using certain receptively known words, even if they may be able to produce them when elicited by the teacher. Laufer and Nation (1999) suggest that "such reluctance is often a result of uncertainty about the word's usage" (p. 37). The present research claims that AV input could be helpful in this aspect, as it provides learners with some information about the lexical items, such as the suitable context where it can be used, its meaning and its pronunciation. This additional information could help them overcome their avoidance of these lexical items.

According to Laufer (2005), another reason for avoiding the use of some known words is the low speed of access, especially in the case of speaking. When time is limited, learners tend to choose words that first come to mind, which leads them to use high frequency words more often, as they can be accessed more quickly. Research has indicated that learners tend to opt for high-frequency words over low-frequency words and more general terms over more specific ones (Laufer, 2005), e.g. *big* over *enormous,* or *looking* over *staring*. The current study seeks to provide insight into whether using AV input might help learners mobilize the less frequent vocabulary that they know receptively. When being exposed to some less frequent

words, learners may be compelled to use them in their own speech as they become more easily accessible.

In conclusion, to encourage learners to use a wider range of words in their free oral production, it may be necessary to focus on two aspects – confidence and speed of access – and to use suitable pedagogical techniques to draw learners' attention to certain lexical items. The next section further explores the possibility of using L2 input for this aim.

2.2   Input and L2 vocabulary development

Input is a crucial component in L2 vocabulary learning and development. Although learners can certainly acquire new vocabulary intentionally through focused vocabulary activities, it has been shown that they can also do so incidentally while engaging with meaningful input. Therefore, learners' attention is focused on the content, rather than the lexical items themselves. Research has found this to happen while reading, listening, and viewing videos (for an overview of studies about different modalities, see Nguyen and Boers, 2019). These studies tend to examine learners' receptive vocabulary knowledge, for example by administering a meaning recall test or asking participants to match the words with their definitions after input exposure. Fewer studies have explored how learners use vocabulary productively in speaking or writing after being exposed to input (for an overview of studies, see Nguyen and Boers, 2019). In most of them the participants are explicitly asked to incorporate specific linguistic items into their oral or written production, thereby shifting their attention and making it a vocabulary-focused activity. The current study, however, seeks to keep focus on meaning, and for this reason, learners are not required to use any specific words. Furthermore, by asking learners to include or use specific words, previous studies have focused on learners' controlled productive vocabulary, as opposed to their free productive vocabulary (i.e. the words learners

would choose to use without any external requirements regarding lexis), which is examined in the present research project.

It may be argued that the studies mentioned above are mostly based on the belief that the instructor or instructional materials should be in charge of choosing the lexical items the learners need to learn or practice. Conversely, some of the more recent studies (Boston, 2008; Hoang and Boers, 2016; Duong *et al.*, 2021) have adopted a learner-centered approach by allowing learners to choose the lexical items they wish to focus on. These studies use the terms *mining of input* and *lexical mining* coined by Samuda (2001), which is understood as the learner's own productive use of linguistic elements taken from input with no prior explicit instruction from the teacher. Since learners are not expressly told to focus on vocabulary, let alone specific target items, it places more focus on learner autonomy by allowing them to decide which items are necessary for the particular task. Giving learners control over selecting words from input may be a useful tool for creating a more learner-centered EFL classroom, though it is necessary to further investigate how successfully learners engage in lexical mining. This is one of the aims of the current thesis.

The term *lexical mining* does not distinguish between words which were previously known at a receptive level and words which were unfamiliar. It merely analyzes to what extent learners incorporate lexical items from the input into their subsequent output during meaning-focused activities. The author of this paper considers this measure very useful for researching productive vocabulary use, as pushing receptively known words into productive use is as important as acquiring new words.

Lexical mining research has examined different modalities of input. One of the earliest studies on lexical mining was carried out by Boston (2008) who investigated the effect of both written and oral input on A2 learners' spoken task performance. He showed that, after being exposed to written task instructions, learners were able to use several linguistic items and

lexical chunks from this input text to complete the subsequent oral tasks. No pedagogical interventions were used to draw learners' attention to any of these items. Surprisingly, the author did not find any lexical mining from oral input (recordings of model dialogues) despite them being played 3-4 times and being highly relevant to the subsequent oral task. The researcher attributed the lack of lexical mining from audio input to the subjects' low proficiency level.

In contrast, Hoang and Boers (2016) found that learners' oral output was heavily influenced by previous video input. The dissimilarity may be due to the participants' higher proficiency level (B1-B2) or differences in the tasks used in each study. In Hoang and Boers' study, learners watched and listened to a story in video format, and then retold it. Therefore, there was an even closer relationship between the input and the learners' output. Additionally, the multimodal nature of the input allowed learners to see the written form of the words in the captions, besides hearing its aural form. This may have facilitated lexical mining.

Hoang and Boers' study (2016) is especially relevant to the current research project, as it introduced a measure to examine the extent of lexical mining, which is also used in this paper. These scholars calculated the number of words occurring in both the input text and the learners' subsequent oral retelling. By comparing it with the number of shared words in the control group who did not receive any input, the researchers were able to gauge the effect of the received input on the learners' productive vocabulary use in the oral task.

Hoang and Boers' results from their 2016 study are consistent with the data obtained by Duong *et al*. (2021) who used a similar measure for lexical mining. They investigated how exposure to short, captioned tourism videos affected the participants' vocabulary use in a narrated imaginary travel experience. These scholars also found that the AV input had a significant effect on learners' subsequent vocabulary use, as they were able to use mined single words and multiword expressions in their follow-up oral performance.

Neither Hoang and Boers (2016) nor Duong *et al*. (2021) identified which mined words were previously known and which ones had been encountered for the first time. Nguyen and Boers (2019), however, focused on specific target words, which were unknown to the participants prior to the experiment, as evidenced by pre-tests. After viewing a TED talk video, the learners were asked to summarize its content. The results indicated that they voluntarily picked up various target words from the input, and immediately used them in their oral retelling. They were also able to retain some of these words in delayed vocabulary post-tests. This shows that not only can learners acquire new lexical items from input, but they can also directly incorporate them into their free productive use. Besides, the video employed in Nguyen and Boers' (2019) investigation was shown without L2 captions, which indicates that seeing the written form of a word is not essential for lexical mining to happen.

The studies presented thus far provide some evidence that learners do reuse lexical items from AV input. However, the findings remain unclear, as different studies have used different designs and have reported considerably different proportions of mined words (e.g. 10% in Duong *et al.* 2021 study vs 68% in Hoang and Boers' 2016 study). The extent of lexical mining may depend on several factors, such as learners' proficiency level, task type, input characteristics, note-taking, salience of lexical items, and their relevance to the subsequent task (Hoang and Boers, 2016). Similarly, individual differences, such as previous vocabulary knowledge and working memory, have been suggested as moderating factors in lexical mining, though studies have not given conclusive results (see studies by Duong, 2021, and Yang, 2017).

Research on lexical mining in written output tasks has similarly given mixed results. El-Dakhs *et al.* (2022) partially replicated Hoang and Boers' 2016 study, using the same input text in a reading-and-listening and a reading-only condition, but asking learners to produce written stories instead of oral narratives. The findings suggest that the input text exerted a strong influence on learners' choice of words, as over 70% of the word families present in the

input also appeared in participants' stories. Conversely, Yang *et al*. (2017) did not find evidence of learners incorporating target words from the reading text into their essays, though this may be attributed to the low relevance of the target words to the essay topic. Hsu (2014) also investigated how input affects vocabulary use in essay-writing. This author showed that using short videos prior to writing may lead to a significant increase in the use of advanced vocabulary, even without any explicit word instruction. Hsu (2014) suggested that input can be useful in two ways: helping learners acquire new unknown words as well as notice words of which they have some partial knowledge, despite not using them productively. This perspective on the usefulness of input is shared by the author of the current thesis.

In conclusion, there is some evidence that input may be a useful pedagogical tool to influence learners' vocabulary use. It seems that the effect of input on learners' vocabulary use can be twofold – on the one hand, learners may incidentally learn new words and immediately use them in their output, and on the other hand, input exposure may push receptively known words into productive use.

2.3   Input processing and noticing

In order to successfully reuse words from input, learners must first notice relevant linguistic items, as simply being exposed to language does not guarantee that it will be attended to or processed (Schmidt, 2010). Schmidt (2010) claims that *noticing*, or focal attention given to a linguistic item, is essential for learning, since "people learn about the things that they pay attention to and do not learn much about the things they do not attend to" (p. 721). While there is not complete agreement on what kind and what extent of noticing is necessary for learning, most scholars agree that "unconscious learning – if it exists at all – is negligible" (Godfroid *et*

al., 2010, p. 173). Therefore, to effectively use input for vocabulary development, it is essential to ensure learners actually notice lexical items in the input they receive.

However, it is not yet entirely clear how exactly learners process input, for example what aspects and features they attend to and to what extent. Measuring learners' noticing is not an easy task, and various methods have their own advantages and disadvantages. Some scholars prefer to collect this data after the task, for instance using questionnaires or interviews so as not to interfere with the task itself (Godfroid *et al*., 2010). However, this creates the challenge of ensuring that learners do not forget what they have noticed. In order to remedy this, some researchers have instructed participants to verbalize or take notes of the aspects they notice while doing the task. More recently, some studies have employed eye-tracking methodology, which registers participants' eye movement while reading, and therefore their attention to specific items (Godfroid *et al*., 2010).

The aim of the current study is not to measure noticing in itself, but rather the incorporation of lexical items, so no special noticing measures were used. While learners were not instructed to verbalize their noticing, several of them made metalinguistic comments during the task. These can offer some insight into their noticing processes and the possible usefulness of output. Similarly, it is assumed that words incorporated into the oral task after input exposure had been noticed. By no means does this study intend to give a comprehensive picture of noticing, as learners may have noticed other items that they chose not to incorporate into their oral production.

Studies on noticing have traditionally focused on whether learners are capable of noticing specific morphological and syntactic forms (for an overview of studies, see Godfroid *et al*., 2010). However, when examining what learners notice in input in general terms, researchers have shown that they overwhelmingly pay attention to lexical items (Williams, 1999). Williams (1999) also claims that this tendency does not depend on the learner's

proficiency level, though the extent to which noticing happens may increase with higher proficiency. An example of this also comes from Hanaoka's (2007) study on noticing in written input, where approximately 65% of all noticing was related to lexis for both intermediate and advanced subjects, though more proficient learners exhibited more noticing overall.

Due to the limited processing capacity of the human brain, not all input can be processed, and learners will need to make a choice about what to consciously process (VanPatten, 1990). In the context of meaning-based activities, such as in the case of the current study, learners will allocate attentional resources to the content and information in the story to comprehend the input. For pedagogical reasons, it is therefore critical to consider whether learners can simultaneously attend to meaning and formal features of the language. Based on their experimental research, VanPatten (1990) suggests that learners struggle to pay attention to both content and grammatical items while processing input. However, this does not seem to be true for language features that carry significant information (i.e. lexical items), which can be noticed and processed without any negative effect on input comprehension (VanPatten, 1990). The researcher claims this to be true for learners at all proficiency levels.

The findings that learners tend to prioritize lexical items and that this does not hinder comprehension of input are clearly advantageous for vocabulary development through meaning-focused input. However, the fact that learners are capable of noticing lexical items, does not necessarily imply that they will do so to a great extent when left to their own devices. Even though input is recognized as beneficial for vocabulary development, it has also been observed that learners often fail to notice linguistic items in input. This has led some researchers to believe that exposing learners to meaningful input without any focus on formal features is not enough for vocabulary development. According to Laufer (2005), the notion that vocabulary development can rely solely on comprehensible input, without incorporating any vocabulary-focused activities, is based on unfounded assumptions. The first assumption is that

learners do indeed pay attention to lexical items in the input. This may often not be the case for several reasons. Firstly, learners tend to overestimate their knowledge of vocabulary in the input, for example they might confuse new words with previously known words (Laufer, 2005). Likewise, it could also be argued that when a word is known receptively, learners do not see the need to focus on it, despite not knowing how to use it in their own speech. Secondly, learners might not attend to unknown words when they are not necessary for understanding the meaning of the input (Izumi, 2003). This may be especially true in the case of AV input where images could provide a great deal of information to aid with comprehension.

Another false assumption is that learners are good at guessing the meaning of new words from the context. This may be difficult as clues to help learners guess may not always be available in the input text or they may be misleading (Laufer, 2005). In this case, AV input may actually be beneficial, especially when new words are represented in the images.

Despite common human tendencies and abilities, there is no doubt that language learners vary greatly, and so does their noticing ability (Schmidt, 2010). Schmidt suggests that various individual differences, such as motivation, language aptitude, and working memory, may influence how much attention language learners pay to linguistic items in the input, though results from empirical studies are inconclusive.

Based on the evidence presented, it appears that noticing and picking up vocabulary from input texts cannot be taken for granted and may be more challenging for some learners than others. Therefore, some type of awareness-raising activities might be beneficial in helping L2 learners focus on vocabulary. Such activities will be discussed in the next section.

2.4 Promoting noticing in input

It seems that incorporating some attention to form in meaning-focused activities may lead to more effective input processing and, by facilitating noticing, may also lead to better learning gains. Several pedagogical interventions have been suggested to increase noticing of linguistic items in input. These include external attention-drawing techniques, such as visual text enhancement (e.g. highlighting some words in a reading text or captions) and explicit instructions from the teacher, as well as internally driven techniques, such as producing output with a view to making learners reflect on their language issues (Izumi 2002). Thus, the main difference between these interventions is who initiates the focus on formal features – the instructor or the learners. Traditionally, the attention given to formal features and lexical items has been controlled by teachers or instructional materials. As Williams (1999) points out, teachers often address learners' needs, by paying more attention to features that the learners are struggling with, or providing words they believe learners might need in the course of a task. However, even with the best intentions, teachers' focus may not always match the learners' needs, and this has sparked a growing interest in learner autonomy in determining the linguistic focus in EFL classrooms.

Similarly, a number of scholars have suggested that learner-driven attention in input processing may be superior compared to teacher-directed attention. Izumi's (2002) experiment on learning L2 relative clauses compared the effectiveness of output-driven noticing (i.e. learner-generated attention) and visual text enhancement by highlighting (i.e. attention generated by instructional materials). The findings evidenced higher learning gains in the former condition, which led Izumi to propose that these two forms of attention are qualitatively different in terms of processing. The researcher suggests that the text enhancement group noticed the highlighted forms in the text but may have not felt the need to process them more

deeply, whereas producing output may have caused the other group to engage in cognitive comparison and more elaborate processing, which facilitated acquisition. This lends support to the use of output tasks as pedagogical tools in EFL classes as opposed to external attention-drawing methods. In a similar vein, Williams (1999) suggests that when focus on formal aspects comes from learners themselves, it indicates their readiness and willingness to acquire them. In other words, learning of new forms may be more effective when learners perceive the immediate necessity to use them, and are then provided with a suitable means of expression. If internally prompted noticing is superior in quality, it is crucial to find out what and to what extent learners notice on their own, to ensure its viability as a pedagogical tool.

Using output as an awareness-raising technique regarding vocabulary has been investigated in several recent experimental studies, especially in combination with input. As mentioned before, when learners process input, they tend to focus mostly on lexical items. The same seems to be true when learners are writing original texts or performing oral tasks, i.e. producing output. Depending on the task type, scholars have reported that lexical issues made up approximately 50-90% of the learner-initiated focus (for an overview of studies, see Hanaoka, 2007). This makes output tasks and their combination with input tasks a particularly suitable technique for vocabulary development.

In their aforementioned studies involving AV input, Nguyen and Boers (2019) and Duong *et al*. (2021) used an output task between two exposures to the same input to draw more attention to lexical items. This type of cycle has been suggested in accordance with Swain's Output Hypothesis, which states that when learners are producing their target language "it may bring their attention to something they need to discover about their second language" (Swain 2005, p. 474). Once learners have realized what they need to discover about their target language, they have several options – they could consult their teacher, peers or a dictionary, try to solve the issue with their own linguistic resources (for instance, paraphrasing), or look

for solutions in relevant input, if available. Nguyen and Boers (2019) advocate for the usefulness of an input-output-input cycle, which would allow learners to revisit the input to address their lexical difficulties:

> 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 (...) to give them the opportunity to fill some of the gaps they noticed in their own resources. (p. 12)

Regarding the first output task, the current study does not propose a summary task of a video, but rather a storytelling task using pictures with no L2 input. This is still believed to help learners notice what linguistic resources they are missing to successfully complete the task, and therefore spark the same noticing function. As the pictures used are the same as the ones presented in the subsequent video input, there is a close relationship between learners' output and the following input. For this reason, when the learners receive AV input after their oral performance, they may be able to find solutions to their lexical issues.

Qi and Lapkin (2001) suggest that when learners fail to solve their linguistic issues with the knowledge they have, this "may result in a sense of uncertainty or lack of fulfillment on the part of a learner" (p. 289). This feeling of lack of fulfillment is believed to lead learners to engage more actively with following input. The fact that learners are eager to find solutions to their problems was also demonstrated by Hanaoka (2007), who asked the participants how interested they were in reading a native-speaker model of the writing task they had just completed. The researcher found that on a scale of 1-5 (1=not interested, 5=very interested), the mean score was 4.3, indicating strong motivation to examine the model texts. Furthermore, Lynch (2018) highlights that, contrary to what many teachers believe, learners actually appreciate the opportunity for a second attempt at their oral task, particularly when they can "engage in some sort of cognitive activity related to their first run" (p. 196). In the current

study, this cognitive activity involves the comparison with a model story, which learners are exposed to after their initial narration.

Empirical studies on output-input cycles with AV input have produced varied outcomes. Nguyen and Boers (2019) found a positive effect of this kind of treatment in their study, where one group summarized the video content between two exposures, whereas the other group did not. The beneficial impact of the output task was reflected in the summary group's higher post-treatment vocabulary test scores. However, it did not seem to have a similar effect in the lexical mining study conducted by Duong *et al.* (2021), who found no significant difference between learners' oral vocabulary use before and after the second viewing. This led them to suggest that while an input-output-input cycle may be beneficial in creating a form-meaning link for new words, it may not be particularly useful for lexical mining. However, in their experiment, the second oral performance was not conducted immediately after the second viewing, but rather two days later, which may have had a negative effect on mining. Furthermore, learners were not aware they would have to repeat the oral task, which may have made them less inclined to find solutions for the lexical issues they had noticed during the first output phase.

The effect of an output-input cycle on noticing and vocabulary use has also been investigated in writing-and-reading tasks. Hanaoka (2007) studied learner-initiated attention in two stages: while composing a story (i.e. the output stage) and in the subsequent input stage, where learners compared their own texts with native-speaker models. The findings suggest that the majority of linguistic features noticed during the comparison stage dealt with problems that had arisen in the initial output phase, for example finding lexical items that the participants had not been able to use on their own. This highlights the usefulness of producing output to promote noticing of items in subsequent input. However, receiving written input also allowed

participants to notice alternative lexical possibilities for words that they had not perceived as problematic originally. Hanaoka (2007) believes that these two types of noticing are distinct:

> These two types of problem recognition may be distinguished from each other. In the case of the former, the learner typically wonders, 'How can I write (say) this?' while in the latter case, the learner may say, 'I should have written (said) it this way.' In this sense, those problems that the participants notice during output, or 'holes' (Swain, 1998) may represent proactive recognition of problems, whereas those problems that they notice for the first time during the comparison stage without being preceded by noticing of 'holes' may be characterized by reactive recognition of problems. (pp. 470-471)

This demonstrates the usefulness of employing such output-input cycles. If learners are producing output and solving their lexical problems with the help of a dictionary or by consulting their teacher, they will only benefit from the first type of problem recognition mentioned by Hanaoka (2007). However, if they compare their own output with relevant input, they may notice further ways to improve their production.

If learners notice gaps in their L2 knowledge and relevant input is then available, it may be useful to consider how successful they are in finding solutions and incorporating them. According to Hanaoka (2007), when participants were given model texts to compare with their own written narratives, they managed to notice two thirds of the lexical solutions that were available in the models and incorporated a vast majority of them in their immediate repeat task. This indicates that when learners notice a gap, they can indeed rather successfully find a solution in input and incorporate those found solutions.

However, writing-and-reading tasks differ considerably from ones that combine speaking and listening due to the higher cognitive demand of the latter. One of the biggest differences is undoubtedly the amount of time available for completing the task. The time constraints in speaking compared to writing tasks may mean that less noticing will occur in the former (Hanaoka, 2006). In a similar vein, the amount of time available for making comparisons with input differs greatly in these two modalities.

Using speaking tasks in combination with AV input, as in the present study, means that learners cannot engage in side-by-side comparisons, which makes the task more challenging and may result in less noticing of gaps and solutions. Similarly, learners may struggle to remember which gaps they had noticed during output, as note-taking is not allowed in the current study, and this might hinder the comparisons with input. However, it has been suggested that when learners notice a gap in their knowledge and cannot find a necessary lexical item, they may remember this aspect more easily. This claim is based on the Zeigarnik effect (Zeigarnik, 1999, quoted from Hanaoka, 2007), which states that incomplete tasks create psychological tension that enhances remembering. Hanaoka (2007) believes that this partly explains why the participants were able to address so many of the gaps they had noticed during output.

Despite some promising results from studies of learner-initiated attention to form in an output-input cycle, it is necessary to bear in mind some problematic factors. Firstly, the answers to learners' issues may not always be present in the subsequent input. This is especially true if there is not a very close relationship between the output task and input, although it can even happen during a story retelling or summarizing task, as there are several ways of interpretation and expression. Another issue is related to individual differences – not all learners are equally capable of noticing gaps in their own knowledge and finding solutions in the subsequent input. Izumi and Bigelow (2000) highlight that some learners may simply not be aware of what linguistic items they need. This became evident in their study on noticing the hypothetical past conditional, where the effect of individual differences was clear both from the test results and participant interviews conducted after the experiment. When asked to produce language in a text reconstruction task, some learners realized that they had difficulties with grammar and therefore looked for solutions in the provided input, while others focused more on ideas and organization of their writing passage. This lends support to the notion that L2 learners show

considerable variability in what they perceive as problematic in their performance. This means that they might focus on different aspects in the subsequent input and may therefore benefit from an output-input cycle to greatly varying degrees. L2 learners' proficiency also seems to play a role in noticing. According to Qi and Lapkin (2001), this holds true for both the noticing of gaps during output and the noticing of solutions in the subsequent input. However, in Hanaoka's 2007 study, there was no significant difference observed between learners of different proficiency levels.

In conclusion, it seems that using an output-input cycle may be useful for helping learners notice their lexical issues while producing language and find solutions in the subsequent input. The benefits include a more learner-centered learning process and possibly higher-quality noticing. In addition, comparing their output with relevant input may allow the learners to find solutions to the lexical issues they had experienced, but also enhance other aspects of their written or oral performance which they had not considered problematic. Learners' overriding focus on lexis makes such treatments especially suitable for vocabulary improvement. However, the role of output in noticing lexical items has not given conclusive results in studies with AV input, and this has prompted the current research project.

2.5   Noticing, output and L2 acquisition

When learners successfully notice lexical items in input and then use them in their output, it may be tempting to claim they have effectively acquired these items. However, this may not be the case, as the relationship between noticing, production and learning is not as straightforward (Alcón, 2007). Examining the effect that noticing and incorporating lexical items has on learners' long-term productive vocabulary use is outside the scope of the current

study, as no delayed oral tasks were carried out. However, some studies that underscore the usefulness of tasks involving noticing and output will be briefly discussed.

Firstly, empirical research has consistently shown that productive practice, i.e. making learners produce the newly learned lexical items, is crucial if the aim is for learners to be able to use these items in speaking and writing in the long run (Laufer, 2005). The same seems to apply to reaching productive mastery of receptively known words (Teng and Xu, 2022). Nguyen *et al*. (2019) also suggest that using an input-output cycle which requires learners to summarize the content of a video allows them to retrieve newly met lexical items from their memory and use them in a generative way, which facilitates vocabulary retention.

The benefits of output were also pointed out by Hanaoka (2007), who found that when learners noticed a gap in their L2 knowledge triggered by output, and thereafter found the missing lexical item in the input, such items were retained better than items mined from input without having originally noticed a gap. This led the researcher to raise the question of whether "the subjective experience of a need for a word increases the impact of noticing a solution and thereby the chance of its retention in memory" (Hanaoka, 2007, p. 475).

Therefore, using tasks that involve output seem to be beneficial for vocabulary acquisition in two ways – firstly by promoting productive mastery of vocabulary, and secondly, by triggering noticing which may lead learners to engage more effectively with lexical items and therefore to retain them longer.

2.6   Audiovisual input in L2 learning

The use of AV materials in the EFL classroom has been gaining popularity over the last decades. They are often considered motivating and engaging for learners and are more readily available thanks to advances in technology. Hsu (2014) also argues that using AV materials

with captions is a good way to accommodate different learning styles, as learners will be able to choose to focus more on written or auditory information, depending on their preferences.

Despite the advantages, Peters *et al.* (2016) caution that vocabulary learning gains from AV input tend to be rather modest. This is likely due to high processing demands compared to written input, and the fact that learners cannot go back to previous words or sentences, at least in classroom settings.

There are however several factors that might foster vocabulary development through AV input. What makes AV input unique is its multimodality – learners can simultaneously attend to visual input (images) and aural input (audio). This means that learners receive information through various channels. Furthermore, if the video has L2 captions or L1 subtitles, they can attend to two types of visual input, instead of one.

One of the most influential theories in the field of AV input is the Dual Coding Theory (Clark and Paivio, 1991). This theory states that humans use two processing systems – verbal (text or audio) and non-verbal (images) – and that these systems can mutually reinforce the acquired knowledge. Muñoz (2022) explains its relevance to L2 learning:

> This theory has been used to explain why L2 learning can be enhanced by combining visual images with verbal information, as when viewing videos. Viewers receive verbal information in the form of aural words, and also of written words in captioned videos, together with non-verbal information in the form of video images. The connections between the two subsystems reinforce one another: videos provide visual referents for the language used, and language can clarify the meaning of an image. (p.126)

Therefore, combining visual, aural and textual information is expected to lead to successful instances of lexical mining.

The role of captions in L2 language learning is still ambiguous, though it seems that captions have a positive effect on both listening comprehension and vocabulary learning. (Peters *et al.* 2016). The combination of on-screen images and text has been suggested to be particularly powerful for several reasons. Firstly, the semantic match between the imagery and the written form in the L2 captions is believed to help create form-meaning links, as visual

clues provide the meaning at the same moment as learners see the form of the word (Peters *et al.*, 2016). According to Muñoz (2022), captions can make language items more salient for L2 learners, and therefore prompt noticing. Similarly, Peters *et al.* (2016) propose that having access to L2 captions enables learners to discern individual words in speech and may therefore facilitate their recognition and acquisition.

While Peters *et al.* (2016) highlight that a higher number of occurrences has a positive effect on learning lexical items from AV input, a meta-analysis by Uchihara *et al.* (2019, quoted from Muñoz, 2022) suggests that frequency plays a smaller role in AV input compared to monomodal input, such as reading and listening. While the current study does not examine whether words with several occurrences were mined more often, it is believed that even items occurring only once might be mined.

Furthermore, recent studies seem to indicate that captions in AV input can even lessen the effect of individual differences, for instance working memory capacity and language learning aptitude (for studies, see Muñoz, 2022), two aspects which previous studies have mentioned as modifying factors in noticing and lexical mining. This gives further support to using L2 captions in the current study, as it may help reduce the impact of individual differences among participants.

In conclusion, using AV input, rather than monomodal input such as reading and listening, should be conducive to lexical mining. Multimodal input is expected to help learners process information through various channels, clarify the meaning of new words, as well as enhance the salience of lexical items.

3   Rationale and Research questions

The evidence reviewed in this paper seems to suggest that AV input may be beneficial to L2 learners' lexical development by influencing their subsequent output in two ways. Firstly, it encourages them to use less frequent vocabulary that they already know receptively. Secondly, it allows them to incorporate newly encountered words. However, much uncertainty still exists about the extent to which learners mine words from AV materials and the factors which may affect that. The current study aims to provide more insight into this. It could help inform teachers' decisions on whether to use AV input for vocabulary development or if other types of interventions would be more justified. Likewise, since relatively little is known about the role of output in helping learners notice useful lexical items in AV input, the present research aims to expand the knowledge in this field. This may provide further information on whether using oral output as an activity for noticing is substantiated. Additionally, it could lend support to giving more space to learner-generation attention in EFL classrooms. Drawing upon the insights gained from the review of the relevant literature, this research aims to address two research questions:

1. To what extent do L2 learners incorporate lexical items from AV input into their subsequent oral output?
2. Does producing output before input exposure promote the incorporation of lexical items in the subsequent oral output?

Regarding the first question, it was hypothesized that learners who were exposed to input would have significantly more shared words with the input text than the learners who did the oral task before receiving input, i.e. the first group would engage in lexical mining. As for the second question, it was expected that the proportion of shared words would be higher for the group

who produced output and was therefore able to see their lexical issues before engaging with input.

To answer these research questions, an empirical study was carried out to collect data on oral vocabulary use in three conditions – with no effect of input, with the effect of input, and with the effect of an output-input cycle. Data was analyzed quantitatively to gauge the extent of lexical mining in different conditions. In addition, a qualitative analysis of learners' metalinguistic comments during their oral performance was carried out to further examine the role of output in noticing.

4   Methodology

4.1   Participants

The participants were 10 L2 learners of English (aged 20-72, 7 females, 3 males) who were native Spanish or Spanish/Catalan speakers. Their English proficiency level was pre-intermediate as they were all students in the one-year B1 course at the state-funded public language schools (Escola Oficial d'Idiomes) in Catalonia, Spain. The experiment was carried out approximately one week before their final exam. While participants were recruited from three different centers, they all followed the same coursebook, English File Intermediate, and the same curriculum set by the Generalitat de Catalunya. They attended 2 in-person classes per week, totaling 130 hours throughout the school year. To be admitted to the B1 group, they were required to have successfully completed an A2 course at the same school or to have passed a placement test for this level.

The learners participated on a voluntary basis but were offered the chance to attend a free speaking workshop after the experiment. Participants were informed that the study was about oral performance and listening skills, but vocabulary skills were not mentioned. For the experiment, they were randomly allocated to one of the two groups: Input-Output Group (n=5) and Output-Input-Output Group (n=5). The Input-Output Group (IO) watched the video and performed the oral task. The Output-Input-Output Group (OIO) performed the oral task twice – before and after watching the video.

4.2   Input

The AV input consisted of a recorded 2-minute story narration that was accompanied by pictures and L2 captions. The input text, a humorous story written by the researcher (see Appendix A), was based on a cartoon (see Appendix B) from the Timesaver Storyboard book. On each of the 8 slides, the participants could see one image, and listen to and read a narration of between 29-44 words. The total length of the story was 279 words.

For both experimental groups, the story was presented simultaneously in three modalities: visual (images), aural (voice-recorded narration), and written (L2 captions). This condition was expected to be conducive to lexical mining for several reasons. Firstly, studies have shown the usefulness of L2 captions in facilitating word form recall, especially for lower-level learners (Peters *et al.,* 2016). Secondly, being able to hear the correct pronunciation might encourage learners to incorporate the words in their oral production, in the case of both previously known and unknown words. Furthermore, the images provided clues to help learners guess the meaning of several possibly unknown lexical items, such as *leash* or *belly*.

The images had been used by the researcher on several occasions with L2 English learners of different levels and were deemed understandable and suitable for B1 level learners. Moreover, it was crucial for the experiment that learners would be able to tell the story just by looking at the pictures – otherwise it would not be possible to collect baseline data (i.e. how they would do the task without any L2 input). Based on previous classroom experience, the images were considered sufficiently self-explanatory to ensure similar retellings of the story among the learners.

The narration was recorded by a British English speaker. The main reason for choosing this variety was its prevalence in the participants' coursebook, and therefore the familiarity with this accent was expected to facilitate aural comprehension.

Regarding the textual input, several aspects were considered when composing the story. As this was a meaning-based activity, it was fundamental that the story was comprehensible for the subjects. Research findings suggest that reasonable comprehension is achieved when learners know between 95-98% of the words in the input (see Kremmel *et al.*, 2023 for an overview of different studies). Similar numbers have been reported across different modalities, namely reading, listening and even TV viewing, with documentaries requiring a lower threshold due to visual aids (Kremmel *et al*., 2023). To evaluate the lexical coverage of the story, the VocabProfile tool was used.[1] This analysis revealed that 2000 most common words achieved the aforementioned necessary lexical coverage of 95%. According to Milton (2010), B1 level English learners typically have a (passive) vocabulary range of 2750-3250 words, which further supports the suitability of the created input.

---

[1] Available at https://www.lextutor.ca/vp/comp/

4.3   Oral task

For the oral task, learners were asked to look at the 8 images one by one and tell the story. They were instructed to say 2-3 sentences about each image. This type of oral retelling task was chosen to ensure a close relation between the input text and oral output. It would create the need to use specific lexical items that are also present in the input and would therefore possibly promote lexical mining. In a similar vein, in the case of the OIO group, the participants would be able to look for solutions for any linguistic issues they had encountered when telling the story from images before receiving input.

4.4   Procedure

The two experimental groups followed the same procedure, with the exception of an added output phase before AV input exposure in the case of the OIO group (see Appendix C for an outline of the procedure).

The participants in the IO group were asked to watch the narrated story and informed that they would have to answer a comprehension question after viewing. In this question, learners had to choose the best title for the story (See Appendix D). This step was included with the aim of ensuring focus on meaning and giving the subjects the opportunity to become familiar with the content of the input. The answers to the comprehension question were not used for analysis. Then, participants were informed that they would be asked to tell the story and instructed to view the narration for a second time. Their oral production was then recorded.

The participants in the OIO Group were first given 2 minutes to look at the 8 images to become familiar with the story before being asked to tell it with the help of the pictures. This oral production was recorded. The experiment proceeded in a similar manner to the IO group.

Participants were asked to watch the video to answer the comprehension question. Subsequently, they were instructed to view the video again. This served as a chance for them to find solutions to the problems they had noticed during the first telling of the story. After having been exposed to the input for the second time, they were asked to retell the story once again while being recorded.

Note-taking was not allowed at any stage of the experiment. Learners were not instructed to pay special attention to vocabulary or any specific linguistic items to ensure that their attention was internally driven, rather than externally imposed.

4.5   Data collection and Analysis

Data was collected in individual face-to-face sessions, and participants' oral production was recorded and later transcribed. For the quantitative data analysis on lexical mining, the transcriptions were trimmed by removing repetitions, filled pauses, metalinguistic comments (e.g. *I don't remember this word*), and phrases not related to the story content (e.g. *In picture number three*). The texts were then lemmatized using the online lemmatizer by the University of Copenhagen.[2] This allowed the author to compare the texts more easily by normalizing inflected forms of the same word. For instance, *lady* and *ladies* are members of the same lemma form, as are *talk*, *talked* and *talking*. Using lemmas, rather than word families, is considered more accurate when researching lower proficiency L2 learners and productive modes (Kyle, 2020). For the qualitative analysis on the role of output, the complete transcriptions were used, with special attention given to the metalinguistic comments which indicated noticing.

---

[2] Available at at https://cst.dk/online/lemmatiser/uk/

As one group performed the oral task once, and the other group twice, there are three different kinds of oral task recordings: 1) before input, 2) after input, and 3) after input with the added effect of previous output. The IO Group produced the oral task once, after viewing the video, therefore their output serves as information about the effect of input on vocabulary use. The OIO Group performed the oral task twice. Their first oral task, completed before viewing the video, served as baseline data to see what words learners at this level of English would use without having received any input (i.e. no effect of input). Their second oral task, completed after viewing the video, provides information about the combined effect of output and input on their vocabulary use. Table 1 provides information about the length of the oral narratives and the number of lemmas in each group. Due to the fact that learners used images to help them with the retelling, there were no big differences in the content of the stories. The only exception occurred in Picture 3 of the story (see Appendix B), where two learners initially thought these people were being supportive, rather than laughing at the runner. Despite this, the oral narratives before and after input are considered to be comparable.

Table 1: Word count and number of lemmas of participants' oral production

	Before input exposure		After input exposure	
	Word count Mean (SD)	Number of lemmas Mean (SD)	Word count Mean (SD)	Number of lemmas Mean (SD)
Group IO (n=5)	-	-	182.6 (60)	45 (11.2)
Group OIO (n=5)	176.6 (46.7)	43.4 (10)	195.2 (28)	48 (8.9)

Word count of input text: 279; Number of lemmas of input text: 84

In order to explore the effect of the AV input on learners' subsequent vocabulary use, this study uses the same measure as previous research on lexical mining (Hoang and Boers, 2016; Duong *et al*., 2021) – the number of shared (content) words between the input text and each of the

recorded oral narratives. Both the input text and participants' oral narratives were lemmatized. The calculation of shared lemmas was carried out using the TextCompare function.[3] The proportion of shared lemmas between the input text and the oral narratives of both groups was then compared to find the effect of input and the effect of the added output phase on participants' lexical choices.

5   Results

5.1   Research question 1: To what extent do L2 learners incorporate lexical items from AV input into their subsequent oral output?

To determine the effect of AV input on learners' subsequent oral vocabulary use, two types of oral narratives were compared: the tasks performed before receiving input (the first oral narratives of the OIO group) and the ones performed after receiving input (the oral narratives of the IO group). Therefore, this is a between-subjects comparison. The number of shared lemmas between the participants' narratives and the original text indicates that being exposed to input had a considerable effect on their choice of words, as shown in Table 2.

Table 2: Number of shared lemmas before and after input exposure

	Number of lemmas shared with input text Mean (SD)
Group OIO Output 1 (Before input exposure) (n=5)	16.8 (1.9)
Group IO Output (After input exposure) (n=5)	24.4 (7.7)

Total number of lemmas in input=84

---

[3] Available at https://www.lextutor.ca/cgi-bin/tl_compare/

Of the lemmas that make up the participants' oral narratives after being exposed to input, on average 24.4 lemmas (29% of the lemmas present in the input text) were shared with the input text, compared to 16.8 shared lemmas (20% of the lemmas present in the input text) in the case of the narratives performed before any input. A two-sample t-test revealed that this difference reached statistical significance, although only at the 0.05 level: $t(4)=2.14$, $p=0.049$.

The fact that learners used about 29% of the lemmas present in the input text does not imply that all these lexical items had been mined from the AV input. It is evident that some of these words would have been used by the learners regardless of the input, as evidenced by the results of the before-input condition, where approximately 20% of the words occurring in the input text were used by the participants. This suggests that these words were already a part of their free productive vocabulary. However, there were several lemmas that none of the learners used prior to input exposure, yet these lemmas appeared both in the input text and the subsequent oral performances.

5.2 Research question 2: Does producing output before input exposure promote the incorporation of lexical items in the subsequent oral output?

In order to explore the effect of an added output phase before input processing, the researcher compared the oral narratives of both groups after having received AV input. While one of the groups had completed an output phase beforehand, the other group had not.

Table 3 shows that the number of shared lemmas between the input text and the immediate oral task was higher for the OIO group than the IO group: 34.2 shared lemmas compared to 24.4. That means that the group who had produced output beforehand used 41% of the lemmas present in the input, as opposed to 29% in the other group.

Table 3: Number of shared lemmas with and without the effect of an added output phase

	Number of lemmas shared with input text Mean (SD)
Group IO Output (After input exposure) (n=5)	24.4 (7.7)
Group OIO Output 2 (After output and input exposure) (n=5)	34.2 (7.2)

Total number of lemmas in input=84

An independent samples t-test indicated a significant difference at the 0.05 level: $t(8)=2.1$, $p=0.036$. Therefore, producing output before engaging with input seems to have a significant positive effect on lexical mining.

6   Discussion

6.1   Quantitative findings

Regarding the first research question, the present study confirms the findings from previous studies on lexical mining from AV input, as learners were able to incorporate words from AV input into their oral production with no explicit vocabulary focus in the instructions. The proportion of shared words falls somewhere in between previous findings – while the current study found that participants used 29% of the lemmas present in the input text, Duong *et al.* (2021) reported 10%, whereas Hoang and Boers (2016) reported 68%. The differences with Duong *et al.*'s study may be attributed to the length of their input text, which had 310 lemmas, compared to 84 lemmas in this MA project. In addition, they used an authentic video, which

may have been more challenging for the pre-intermediate learners, and the output task was not so closely related to the input video.

The differences with the results reported by Hoang and Boers might stem from different task conditions. While the input and output tasks were similar (an oral retelling of a narrated story with images and L2 captions), these scholars allowed the participants to take notes while viewing the video, which may have facilitated the use of words from the input. Besides, the participants in that study were slightly more proficient.

Regarding the second research question, an output-input cycle enhanced lexical mining, as learners who had told the story before watching the video were able to incorporate significantly more words from the AV input than the other group. Therefore, the input text appears to have a strong effect on the learners' choice of words in their repeat oral task.

It seems that the output task served to help learners *notice gaps* (Swain, 2005) in their knowledge and look for solutions in the input. This may explain the higher proportion of mined words. Further evidence of the effectiveness of the output-input cycle comes from qualitative analysis, which will be presented in the following section.

Since no studies to the author's knowledge have used an output-input-output cycle to investigate its effect on lexical mining, comparison with previous research is difficult. Duong *et al.* (2021), who used an input-output-input cycle with a delayed output phase two days later, did not find a positive effect of the initial output on lexical mining. This may be explained by differences in study design.

The results of the current investigation are consistent with the study by Nguyen *et al.* (2019), who found that producing output helped learners notice more target items from the input than simply answering comprehension questions. However, they looked at specific target items unknown to the learners and did not measure lexical mining in general.

Output-input-output cycles have been used in experiments with written output (Qi and Lapkin, 2001; Hanaoka, 2007), though they focused more on the nature of noticing and the incorporation of solutions, rather than measuring the extent of lexical mining. Nevertheless, the positive effect of output on noticing gaps and solutions was evident for both lexical and grammatical aspects in these studies.

6.2 Qualitative findings

To further investigate the quantitative findings, the participants' oral narratives were examined with particular emphasis placed on their metalinguistic comments where the learners spoke about a linguistic problem they encountered. Additionally, attention was paid to pauses occurring in the middle of a clause, as this usually signals that learners are trying to retrieve a word from their memory (Nguyen *et al*., 2019). These aspects provided additional insight into the type of noticing which occurred during the output and input phases, as well as its potential impact on the vocabulary in their repeat oral task.

In accordance with previous research (Williams, 1999; Hanaoka, 2007), the participants in this study seemed to focus mostly on lexis. Evidence of this comes from several metalinguistic comments during the oral task, such as *I don't know the word I want to say* and *I don't remember exactly the word*, as well as learners' comments after the experiment. No comments appeared to be explicitly about grammar, though some comments were vague, for instance *no sé si lo he dicho bien* (Spanish for *I don't know if I said it well*).

As suggested by the metalinguistic comments and pauses, learners noticed gaps in their knowledge when producing output. When processing input, they were sometimes able to find solutions to some of their lexical issues. They also noticed other useful items that were then

incorporated into their repeat oral task. This is in line with Hanaoka's (2007) suggestion that when comparing their own production with a model, two types of noticing might take place – when learners see how to say a word they did not know before, and when they realize they could have expressed something in another (better) way. The first type of noticing is visible when comparing a learner's gaps in the first output task with solutions in the second. The second type, however, may not be apparent in the first output, but can be seen in comparison with the second narration after having received input. The following sections analyze some examples of each, and additionally highlight some issues with using an output-input cycle for vocabulary improvement.

6.2.1   Lexical mining: Noticing the gap

Several evident instances were observed where subjects from the OIO Group struggled to find appropriate words during the initial output task (telling the story from pictures). Therefore, they had *noticed the gap*, according to Swain's (2005) definition. As they were then exposed to the video input, they had the opportunity to try to find the lexical items they needed.

The following examples present some initial output before input exposure (Output 1), relevant input, and the repeat output after input exposure (Output 2). Metalinguistic comments and L1 words are marked in italics, whereas longer pauses are marked with three dots. Participants' names have been changed to protect their anonymity.

A successful case of lexical mining can be seen in (1) – the learner expresses their lack of knowledge of the word *bite* but uses it successfully after watching the video. Notably, immediately after completing their initial oral performance, the learner requested this specific word (using their L1) from the researcher, unaware they would soon receive input featuring the

same story. As the item was not provided by the researcher, it is reasonable to believe that the learner was focused on locating it in the input.

> (1)
> **Carme Output 1:** He is injured and the dog *no sé I don't know the word* him.
> **Input text:** One of the ladies got furious and started shouting at Tony. The dog was also furious and bit Tony's arm.
> **Carme Output 2:** The dog bit him at him in the arm and the lady, a lady *I don't know.*

In the next example, the participant is searching for the English translation of the L1 word *cansado* (Spanish for *tired*), and resorts to using the word *wasted* on three occasions thereafter. However, after viewing the video, he only produces the word *exhausted*, which is present in the input text, and was therefore possibly mined from the input. Similarly, the word *leash* seems to have been mined from the input as it was previously unknown to the learner.

> (2)
> **David Output 1:** The man I think that he didn't train before and he seems *I don't know cansado I don't know the word now*.
> **David Output 1**: In the picture four I can see the same man that it's so wasted.
> **David Output 1**: The man touch the… touch the dog *I don't know the name of*… and he falls to the street and the dog too.
> **Input text:** But he was so exhausted that he didn't see their small dog on a leash.
> **David Output 2**: But Tony was so excited exhausted and he didn't see the leash of the dog.

This aligns with findings by Nguyen and Boers (2019) who showed that learners can and do incorporate newly met words from the input into their subsequent oral summary task. Asking them to retell the content of the video creates a "near-immediate opportunity (...) to try using newly met words" (Nguyen and Boers, 2019, p. 21).

In (3), the learner realized that the word they had wanted to use was not *dry* but failed to recall the word *wet*. Arguably, this word is known to them, judging by their metalinguistic comment. However, after receiving input and noticing the relevant word, the learner has found a solution to their issue, as the word became more easily accessible. This demonstrates the importance of being able to access lexical items quickly during speaking activities (Laufer, 2005), and how AV input could help with that.

(3)
David Output 1: So in the last picture, the eight, I can see the man, it's so dry dry *seco I don't know now*
Input text: Tony got back home, exhausted, disappointed and completely wet.
David Output 2: When he arrived to the house he arrived too wet.

Being able to compare their initial output to a model seems to have helped the learners choose more accurate words and lexical forms, after having resorted to an unsatisfactory word in the first oral task. In (4), the learner paused before using the word *book*, signaling difficulties finding a suitable lexical item to describe the picture.

(4)
Adrià Output 1: One day he saw in a...in a book some advertising from marathon race.
Input text: Later that morning, he was reading a fitness magazine when he noticed an advertisement for a marathon in his city.
Adrià Output 2: A little later in this morning he saw in a running magazine an advertisement who was about marathon race in his city.

Altogether, these examples indicate that learners experienced some difficulties choosing the best words to express their ideas during the initial output phase. Their use of pauses and metalinguistic comments allows for examining the instances where noticing of linguistic problems happened during output, and whether the participants were able to find solutions in the input. The successful cases of noticing the gap and, more importantly, noticing the solutions, illustrate the usefulness of an output-input cycle. When learners were given the opportunity to receive input, they used it to mine the input for necessary lexical items.

6.2.2 Lexical mining: Lack of noticing the gap

On several occasions there were no signs that the learner had noticed a gap or was struggling to find a suitable word (no pauses or metalinguistic comments) but yet they engaged in lexical mining. Therefore, it is possible that the learner had not noticed anything problematic in their output or had simply decided to avoid the problematic item. However, on hearing how the same idea was expressed in the model, they decided to change their original choice of words. (5)

shows how the learner changed their use of words *fat* and *sad* and opted for expressions *gained ten kilos* and *disappointed* in their narrative after having received input.

(5)
David Output 1: He can see herself so fat and she's *sorry* he's sad with this image.
Input text: One morning, Tony was looking at himself in the mirror, and he realised that he had gained about ten kilos. He was shocked and disappointed to see how big his belly was.
David Output 2: One morning Tony the name, see that he gained ten kilos and he was disappointed with this image.

This demonstrates the value of using models as input to help learners improve their lexis. It is likely that if the learner has not noticed a gap in their knowledge and has already found a satisfactory lexical item to express their ideas, they will not consult their teacher or a dictionary. On the other hand, hearing the word in the input text may prompt them to use a wider range of less frequent vocabulary and more accurate words.

Several examples were found where participants swapped more general or high-frequency words in their first oral task for more specific or lower-frequency counterparts in their repeat oral task, probably due to lexical mining. Some of these changes include *tired* for *exhausted*, *talking* for *chatting*, *angry* for *furious*, *advertising* for *advertisement*, *newspaper* for *fitness magazine, wet* for *completely wet*. While these changes were not necessary to retell the story, learners saw a chance to upgrade their language. According to Hoang and Boers (2016), learners do not exclusively mine words that are essential to the task and vital to the story. Similar evidence emerges in the current data, where the most often mined lexical item was *a group of*, which could be easily omitted or substituted. Surprisingly, it was present in eight retellings after video-viewing, despite not being used by any participants prior to input.

6.2.3 Issues using an output-input cycle

However, as mentioned in the Sections 2.3 and 2.4, using input for vocabulary development has its own challenges, whether used in combination with output or not. Firstly, learners were not always able to remember the words they had seen in the input, despite some evidence that they had noticed them while viewing the video. In (6), the participant has clearly noticed *trainers* and *shorts* in the input, as they were not present in their initial narration. During their second narration, though, they were only capable of producing the word *shorts* and failed to find L2 word for *zapatos* (Spanish for *shoes*).

> (6)
> **Carme Output 1:** He began to running at on the street but the neighbours laughed at him.
> **Input text**: The next day, he put on his trainers and shorts and went out for a run.
> **Carme Output 2**: Then he decided to get fit and in the morning he get out to run with shorts and *I don't remember zapatos.*

Difficulties recalling words from the input were also evident in the IO group. In (7), the learner is trying to remember the word *leash*. As they watched the video twice before retelling the story, and the word *leash* appeared 3 times in the story, it seems that even 6 encounters may not be enough to be able to mine some items. Interestingly, this participant asked the researcher to provide this L2 word right after the experiment had finished. This illustrates the lack of fulfillment learners may feel when they fail to find a lexical item (Qi and Lapkin, 2001), and how this feeling may lead to learners' eagerness to find solutions to their linguistic problems (Hanaoka, 2007).

> (7)
> **Ainhoa Output**: He didn't saw the *I don't remember the name the lat* no I don't remember the word* but he didn't see this…this part of the… *I don't know I don't remember* but he don't see that and he fell into the floor, and get the dog with throw away with him.

These instances highlight some of the possible disadvantages of AV input. As discussed in Section 2, listening-and-speaking tasks have a high cognitive demand (Hanaoka, 2006), and

they do not allow the learner to return to the word in the input while they are performing the oral task. Therefore, such problems are understandable, as the learner has to retrieve the item from their memory. Perhaps having learners produce output before input exposure could help with this, since their attention would be more focused on finding solutions to the lexical issues they had encountered.

Similarly, there were instances where learners made unsuccessful attempts to incorporate some lexical items from the input, resulting in incorrect form. In (8) the participant successfully mined the unknown word *leash*, but had difficulties with the expression *tripped over*, producing *pulled over* instead. Multiword expressions such as phrasal verbs are known to be especially difficult to mine from input (Hoang and Boers, 2016).

> (8)
> **David Output 1**: The man touch the… touch the dog *I don't know the name of*… and he falls to the street and the dog too.
> **Input text**: Tony did not notice the dog or the dog's leash, so he tripped over it and fell on the ground. The tiny dog flew through the air, pulled by its leash.
> **David Output 2:** He didn't see the leash of the dog and he pulled over *I think so* pulled over and the dog is flew at the air.

However, the fact that the learner was unable to use the word or to use it correctly, does not mean that noticing was not beneficial. In the case of newly met words, the learner may have only gained some partial knowledge, but not enough to be able to use it productively.

Furthermore, the learners in the OIO group were not always successful in finding solutions to their lexical issues. This may have been due to the fact that the learner was not able to notice/understand the word, despite it being available in the input. (9) shows the learner's difficulties in expressing their ideas in English, as indicated by the metalinguistic comments and the use of L1 words. Despite noticing the gap, the learner did not succeed in mining words from the input, but instead tried to change the L1 word *chocar* (Spanish for *crash into*) into an English word *chock**.

(9)
Estela Output 1: Because it's too much tired and can't see nothing, and she *I don't know in English tropieza.* We find a dog and can't see and fall in the street *I don't know if I say correctly because I don't know the word.*
Input text: But he was so exhausted that he didn't see their small dog on a leash. Tony did not notice the dog or the dog's leash, so he tripped over it and fell on the ground.
Estela Output 2: He was too much exhausted for…put attention and coordinate his feet. And he find found later a two ladies and he didn't see the little dog and… *chock** with the [laughs] *chock** with the dog and fall to the road.

Sometimes, the solutions were not available in the video input. This has been highlighted as an issue when using L2 models to help learners with their linguistic gaps (Hanaoka, 2006). (10) illustrates how the participant is trying to recall the L2 translation for the L1 word *tard* (Catalan for *late*). This L2 word was not available in the input, making lexical mining impossible. In their second narration, the learner omitted this piece of information altogether.

(10)
Estela Output 1: The man is horrible day because she arrived at home very tired and very…*uhm*… very *uhm…tard* very *I don't remember...tard I don't remember*…tired and *uhm I don't remember the*…and in the night, tired in the night

It is also likely that on some occasions lexical mining did not occur because learners were satisfied with their choice of words, even if it differed from the original text. After all, they had been instructed to retell the story, with no specific instructions to try to reproduce it as closely to the original as possible. Therefore, the learner may have decided not to mine words from the video. In (11), we can see that the subject expressed himself very similarly in their repeat task, despite differences with the received input. This indicates that even tasks with a close relationship between input and output may fail to promote mining of some lexical items. Explicit instructions to replicate the story as faithfully as possible could help in this regard, if deemed necessary.

(11)
Adrià Output 1: Suddenly that dog is the is a problem to the runner because he cross in front of him and he fell on the floor.
Input text: But he was so exhausted that he didn't see their small dog on a leash. Tony did not notice the dog or the dog's leash, so he tripped over it and fell on the ground.
Adrià Output 2: He was very exhausted and he fell on the ground because the dog crossed in front of him.

In (12), not only did the learner stay true to their initial choice of words, but even decided to choose a different name for the character (*John* as opposed to *Tony*). This suggests that they attempted to tell their own version of the story, rather than a simple copy.

> (12)
> **Elena Output 1:** I can see middle-aged man who looks upset about himself. I think he looks overweight and he is thinking about do he has to do something about it.
> **Input text:** Tony was looking at himself in the mirror and he realised that he had gained about ten kilos. He was shocked and disappointed to see how big his belly was.
> **Elena Output 2:** One day in the morning, I call this character John, was looking at the mirror and he looks overweight, he looks himself overweight.

The fact that learners varied greatly in the extent of lexical mining is also evident when comparing the number of shared lemmas between the initial output and the repeat output after exposure to input. The number of lemmas that learners picked up from the AV input ranged from 8 to 28, indicating that some learners have a greater tendency and/or ability to engage in lexical mining.

In conclusion, there seems to be some evidence that mined words included both previously unknown words as well as receptively known words, which may not have reached productive use or were simply not accessible enough. In addition, the qualitative analysis provides further support to the noticing function of output, according to Swain's Output Hypothesis (2005), as the output phase brought learners' attention to gaps in their linguistic knowledge. Upon receiving input, they were able to fill some of these gaps. Learners in the OIO group engaged in more lexical mining, which indicates that they were more successful at noticing the lexical items they needed.

In addition, using models from proficient speakers as input might help learners enhance their vocabulary use beyond simply filling the linguistic gaps they had noticed while attempting the task themselves. This may be especially true for more specific or lower-frequency words that learners have some knowledge of, but do not normally use productively. Seeing and hearing them in AV input could promote their use.

However, it must be kept in mind that using output-input-output cycles has some possible drawbacks. For instance, learners may not always succeed in finding answers to their lexical problems and, in general, some learners may have a lower tendency to mine lexical items from the input or more difficulties in doing so. Furthermore, lexical mining from videos may be challenging for some learners due to the cognitively demanding nature of this type of input.

7   Conclusion

This empirical study indicated that receiving AV input has a significant influence on learners' subsequent oral vocabulary use, as learners noticed lexical items and incorporated them into their oral production. This means that by using input and output tasks that are closely related, teachers can help learners move beyond the words they would normally use and help them employ a wider range of vocabulary in their oral performance, even without specific word-focused activities. Furthermore, asking learners to engage in oral output tasks before video exposure may help them notice and incorporate even more lexical items from the input.

These findings suggest that using AV input could be an interesting pedagogical tool for vocabulary development. This kind of task sequence, which combines input and output could make EFL classrooms more learner-centered, as each learner is free to choose linguistic items from the input according to their needs and wishes. However, the success of lexical mining may be subject to individual variability, and therefore instructors must use their own judgment and knowledge of their learners when deciding whether to use this type of intervention. Furthermore, its viability with authentic input materials and different kinds of tasks still needs to be empirically verified.

This study has several limitations. Due to the small number of participants, the generalizability of the findings may be limited. Perhaps more importantly, the current study

used input materials specifically created for the experiment, whereas authentic materials are favored in modern EFL classrooms. Using authentic video input in a similar study design would be an opportunity for further research. Other possible research directions include empirical studies with learners of different proficiency levels, different oral task types (i.e. not only retelling tasks), other linguistic forms (e.g. grammatical and morphological features), and the possible effectiveness of note-taking and explicit instructions to focus on vocabulary. In addition, the current study did not examine the long-term effects of this treatment on the learners' productive vocabulary. It would be fruitful to explore the learners' oral vocabulary use in a delayed task to see whether the OIO group retained their advantage.

Appendices

Appendix A: Input text

One morning, Tony was looking at himself in the mirror and he realised that he had gained about ten kilos. He was shocked and disappointed to see how big his belly was.

Later that morning, he was reading a fitness magazine when he noticed an advertisement for a marathon in his city. So he decided to get fit and take part in this marathon. He felt really excited about his new goal.

The next day, he put on his trainers and shorts and went out for a run. Unfortunately, he was not very fit and quickly started to feel exhausted. He was sweating a lot and a group of teenagers were pointing and laughing at him.

Tony ignored the group of boys and continued running. When he was next to the park, he saw two ladies chatting in the street. But he was so exhausted that he didn't see their small dog on a leash.

Tony did not notice the dog or the dog's leash, so he tripped over it and fell on the ground. The tiny dog flew through the air, pulled by its leash.

One of the ladies got furious and started shouting at Tony. The dog was also furious and bit Tony's arm. It all happened so fast that he didn't have time to react.

After this incident, Tony was walking home and, to make the situation worse, it started to rain! His arm was hurting a lot and his shoes were getting wet in the rain.

Tony got back home, exhausted, disappointed and completely wet. He took the fitness magazine and threw it in the bin. Running a marathon - what a silly idea!, thought Tony.

Content words (84)
advertise air arm belly big bin bite boy chat city complete continue day decide disappoint dog excite exhaust fall fast feel fit fly fortunate furious gain goal ground group happen home hurt idea ignore incident kilo lady late laugh leash look magazine make marathon mirror morning next new notice park part point pull put quick rain react read realize really run see shock shoe shorts shout silly situation small start street sweat take teenage think throw time tiny train trip two walk wet worse

Appendix B: Images for the AV input
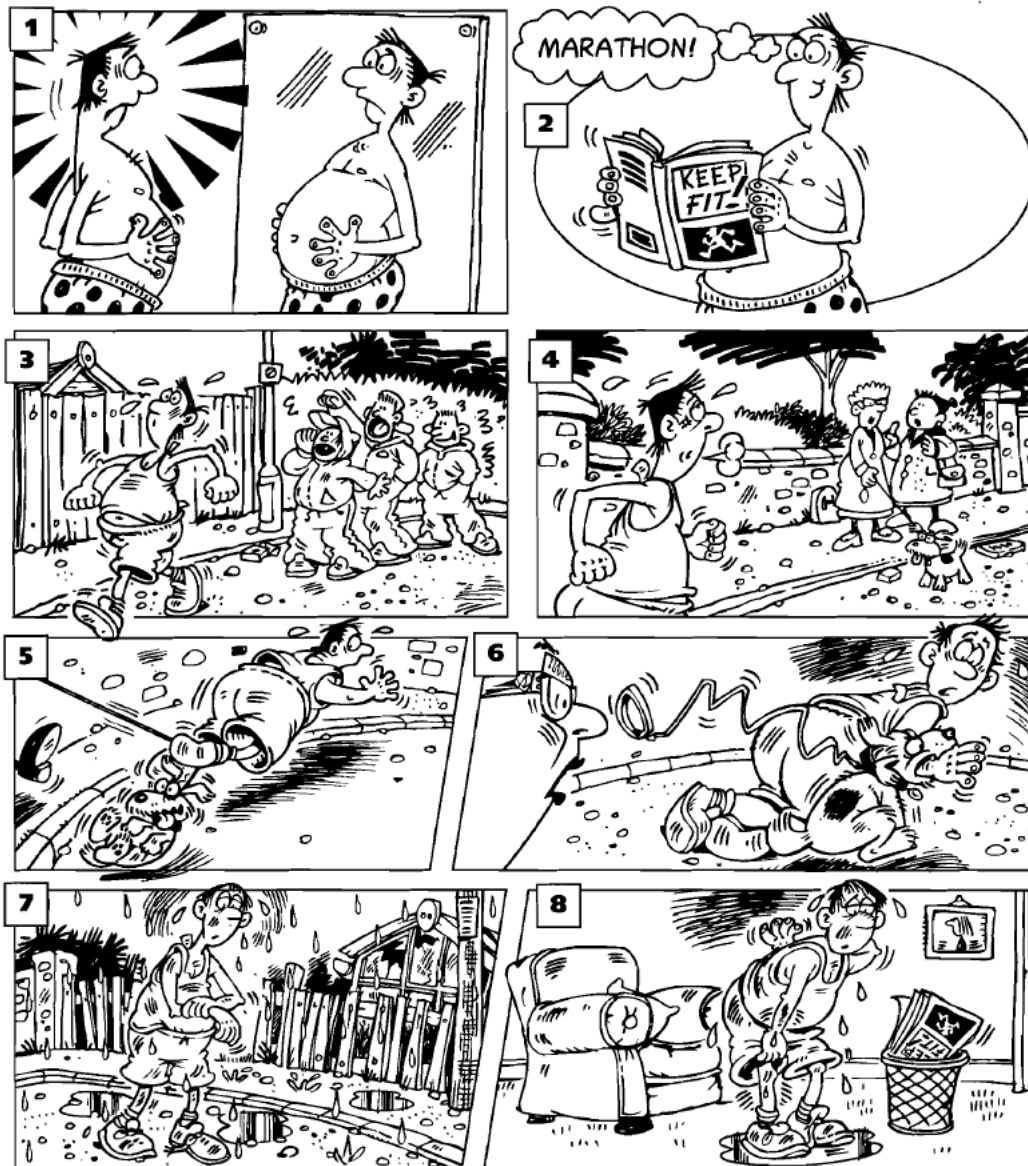
Appendix C: Experiment outline

Group IO	Group OIO
	Images ↓
	OUTPUT 1: tell the story (oral task without the effect of input) ↓
AV INPUT: images+audio+captions ↓	AV INPUT: images+audio+captions ↓
Comprehension question ↓	Comprehension question ↓
AV INPUT: images+audio+captions ↓	AV INPUT: images+audio+captions ↓
OUTPUT: retell the story (oral task with the effect of input)	OUTPUT 2: retell the story (oral task with the effect of output and input)

Appendix D: Comprehension question

In your opinion, which of these 3 titles is the best for this story?

1. Tony's bad luck
2. The winner and the loser
3. The finish line

References

Alcón, Eva. (2009). Incidental focus on form, noticing and vocabulary learning in the EFL classroom. *International Journal of English Studies*, *7*(2), 41–60. Retrieved from https://revistas.um.es/ijes/article/view/48981

Boston, Jeremy Scott. (2008). Learner mining of pre-task and task input. *ELT Journal*, *62*(1), 66–76. https://doi.org/10.1093/elt/ccm079

Clark, James M. & Allan Paivio. (1991). Dual coding theory and education. *Educational Psychology Review*, *3*(3), 149–210. https://doi.org/10.1007/bf01320076

Duong, Phuong Thao, Maribel Montero Perez, Long Quoc Nguyen, Piet Desmet & Elke Peters. (2021). Incidental lexical mining in task repetition: The role of input, input repetition and individual differences. *System*, *103*, 102650. https://doi.org/10.1016/j.system.2021.102650

El-Dakhs, Dina Abdel Salam, Shazia Khalid Khan & Maram Al-Khodair. (2022). Do foreign language learners mine input texts for multiword expressions? The case of writing story retellings. *Ampersand*, *9*, 100080. https://doi.org/10.1016/j.amper.2021.100080

Fletcher, Mark & Richard Munns. (2005). *Timesaver storyboard: 24 stories through pictures*. Scholastic.

Godfroid, Aline, Alex Housen & Frank Boers. (2010). 10. A procedure for testing the Noticing Hypothesis in the context of vocabulary acquisition. In M. Pütz & L. Sicola (Eds.), *Converging Evidence in Language and Communication Research* (Vol. 13, pp. 169–197). John Benjamins Publishing Company. https://doi.org/10.1075/celcr.13.14god

Hanaoka, Osamu. (2006). Exploring the role of models in promoting noticing in L2 writing. JACET Bulletin. 42. 1-13.

Hanaoka, Osamu. (2007). Output, noticing, and learning: An investigation into the role of spontaneous attention to form in a four-stage writing task. *Language Teaching Research*, *11*(4), 459–479. https://doi.org/10.1177/1362168807080963

Hoang, Hien & Frank Boers. (2016). Re-telling a story in a second language: How well do adult learners mine an input text for multiword expressions? *Studies in Second Language Learning and Teaching*, *6*(3), 513–535. https://doi.org/10.14746/ssllt.2016.6.3.7

Hsu, Wenhua. (2014). The effects of audiovisual support on EFL learners' productive vocabulary. *ReCALL*, *26*(1), 62–79. https://doi.org/10.1017/S0958344013000220

Izumi, Shinichi. (2002). Output, input enhancement, and the noticing hypothesis: An Experimental Study on ESL Relativization. *Studies in Second Language Acquisition*, *24*(4), 541–577. https://doi.org/10.1017/S0272263102004023

Izumi, Shinichi. (2003). Comprehension and production processes in second language learning: In search of the psycholinguistic rationale of the output hypothesis. *Applied Linguistics*, *24*(2), 168–196. https://doi.org/10.1093/applin/24.2.168

Izumi, Shinichi & Martha Bigelow. (2000). Does output promote noticing and second language acquisition? *TESOL Quarterly*, *34*(2), 239. https://doi.org/10.2307/3587952

Kremmel, Benjamin, Bimali Indrarathne, Judit Kormos & Shungo Suzuki. (2023). Unknown vocabulary density and reading comprehension: replicating Hu and Nation (2000). *Language Learning*, *73*(4), 1127–1163. https://doi.org/10.1111/lang.12622

Kyle, Kristopher. (2020). Measuring lexical richness. In Stuart Webb (Ed.) *The Routledge Handbook of Vocabulary Studies*, 454–476. https://doi.org/10.4324/9780429291586-29

Laufer, Batia. (2005). Focus on form in second language vocabulary learning. *EUROSLA Yearbook*, *5*, 223–250. https://doi.org/10.1075/eurosla.5.11lau

Laufer, Batia & Paul Nation. (1999). A vocabulary-size test of controlled productive ability. *Language Testing*, *16*(1), 36–55. https://doi.org/10.1191/026553299672614616

Laufer, Batia, & T. Sima Paribakht. (1998). The relationship between passive and active vocabularies: Effects of language learning context. *Language Learning*, *48*(3), 365–391. https://doi.org/10.1111/0023-8333.00046

Lynch, Tony. (2018). Chapter 8. Perform, reflect, recycle: Enhancing task repetition in second language speaking classes. In M. Bygate (Ed.), *Task-Based Language Teaching* (Vol. 11, pp. 193–222). John Benjamins Publishing Company. https://doi.org/10.1075/tblt.11.08lyn

Milton, James. (2010). The development of vocabulary breadth across the CEFR levels. In I. Vedder, I. Bartning, & M. Martin (Eds.), *Communicative proficiency and linguistic development: intersections between SLA and language testing research* (pp. 211- 232). European Second Language Association. https://eurosla.org/monographs/EM01/211-232Milton.pdf

Muñoz, Carmen. (2022). Audiovisual input in L2 learning. *Language, Interaction and Acquisition*, *13*(1), 125–143. https://doi.org/10.1075/lia.22001.mun

Nguyen, Chi-Duc & Frank Boers. (2019). The Effect of content retelling on vocabulary uptake from a TED talk. *TESOL Quarterly*, *53*(1), 5–29. https://doi.org/10.1002/tesq.441

Peters, Elke, Eva Heynen & Eva Puimège. (2016). Learning vocabulary through audiovisual input: The differential effect of L1 subtitles and captions. *System*, 63, 134–148. https://doi.org/10.1016/j.system.2016.10.002

Qi, Donald S., & Sharon Lapkin. (2001). Exploring the role of noticing in a three-stage second language writing task. *Journal of Second Language Writing*, *10*(4), 277–303. https://doi.org/10.1016/S1060-3743(01)00046-7

Qian, David. D. & Linda H. Lin. (2019). The relationship between vocabulary knowledge and language proficiency. In Stuart Webb (Ed.) *The Routledge Handbook of Vocabulary Studies*, 66–80. https://doi.org/10.4324/9780429291586-5

Samuda, Virginia. (2001). Guiding relationships between form and meaning during task performance: the role of the teacher. In M. Bygate, P. Skehan, and M. Swain (Eds.) *Researching Pedagogic Tasks : Second Language Learning, Teaching and Testing.* Longman

Schmidt, Richard. (2010). Attention, awareness, and individual differences in language learning. In W. M. Chan, S. Chi, K. N. Cin, J. Istanto, M. Nagami, J. W. Sew, T. Suthiwan, & I. Walker (Eds.) *Proceedings of CLaSIC 2010*, December, 721-737. National University of Singapore, Centre for Language Studies.

Schmitt, Norbert. (2014). Size and depth of vocabulary knowledge: What the research shows. *Language Learning*, *64*(4), 913–951. https://doi.org/10.1111/lang.12077

Schmitt, Norbert. (2019). Understanding vocabulary acquisition, instruction, and assessment: A research agenda. *Language Teaching*, *52*(2), 261–274. https://doi.org/10.1017/S0261444819000053

Swain, Merrill. (2005). The output hypothesis: Theory and research. In Hinkel, E. (Ed.) *Handbook of research in second language teaching and learning*. (pp. 471-483). Routledge https://doi.org/10.4324/9781410612700

Teng, Min, & Jinfen Xu. (2022). Pushing vocabulary knowledge from receptive to productive mastery: Effects of task type and repetition frequency. *Language Teaching Research*. https://doi.org/10.1177/13621688221077028

VanPatten, Bill. (1990). Attending to Form and Content in the Input: An Experiment in Consciousness. *Studies in Second Language Acquisition*, *12*(3), 287–301. https://doi.org/10.1017/S0272263100009177

Webb, Stuart. (2008). Receptive and productive vocabulary sizes of L2 learners. *Studies in Second Language Acquisition*, *30*(01), 79-95. https://doi.org/10.1017/S0272263108080042

Williams, Jessica. (1999). Learner-generated attention to form. *Language Learning*, *49*(4), 583–625. https://doi.org/10.1111/0023-8333.00103

Yang, Yingli, Natsuko Shintani, Shaofeng Li & Yingyi Zhang. (2017). The effectiveness of post-reading word-focused activities and their associations with working memory. *System*, *70*, 38–49. https://doi.org/10.1016/j.system.2017.09.012