**Article:**
Dang, TNY orcid.org/0000-0002-3189-7776, Lu, C and Webb, S (2022) Incidental Learning of Collocations in an Academic Lecture Through Different Input Modes. Language Learning, 72 (3). pp. 728-764. ISSN 0023-8333

https://doi.org/10.1111/lang.12499

**Incidental learning of collocations in an academic lecture through different input modes**

**Thi Ngoc Yen Dang**

University of Leeds, U.K

**Cailing Lu**

Shanghai Jiao Tong University, China

**Stuart Webb**

University of Western Ontario, Canada

## Abstract

In this quasi-experimental study, 165 EAP learners at a university in China were randomly assigned to five experimental groups and a control group. Each experimental group encountered 19 target collocations in the same academic lecture in one of the following input modes: (a) reading, (b) listening, (c) reading while listening, (d) viewing, and (e) viewing with captions. The control group did not receive any treatment. The results revealed that reading, viewing, and viewing with captions led to learning at the form recognition level, but no significant differences were found in the learning gains across these modes. Nonverbal elaboration, type of vocabulary, and type of verbal elaboration affected learning, but frequency of occurrence, strength of association, comprehension, and prior knowledge of general vocabulary did not. This study provides further evidence supporting the use of academic lectures for incidental learning of collocations as well as expanding on the multimedia learning theory.

***Key words:*** *incidental vocabulary learning; academic lecture; multimedia learning theory; viewing; collocation*

## Introduction

Knowledge of collocations is essential for high levels of second language (L2) proficiency (Nation, 2013). However, mastering collocations is challenging for many L2 learners (Nguyen & Webb, 2017). While collocations can be learned through explicit learning activities, it would be impossible for teachers to explicitly teach all collocations to learners (Puimège & Peters, 2020). Therefore, apart from explicit instruction, it is important to identify sources of input for learners to acquire collocations incidentally. Studies investigating incidental learning of collocations are limited in number and have only focused on non-academic sources of input such as graded readers (e.g., Pellicer-Sánchez, 2017; Webb & Chang, 2020), songs (Pavia, Webb, & Faez, 2019), and television programs (e.g., Puimège & Peters, 2019, 2020; Teng, 2019). Recently Dang, Lu, and Webb (2021) found that collocations could be learned through viewing a single academic lecture. This suggests that academic lectures may be a potential source for incidental learning of collocations. With the increasing popularity of open access courses delivered by well-known universities, academic lectures are now freely available in different formats (transcripts, audios, videos, and videos with captions). This means that if academic lectures from these sources are used as meaning-focused input for L2 learners, they can be encountered in various modes: reading, listening, reading while listening, viewing, and viewing with captions. To help EAP teachers and learners make better use of this kind of L2 input, it is essential to determine which of these modes best facilitates incidental learning of collocations in academic lectures. On the one hand, Webb and Chang's (2020) research on incidental learning of collocations in graded

readers revealed that reading and listening led to a fairly similar amount of learning gains. This finding suggests that if academic lectures were used as L2 input, there would probably be no significant difference in the collocational learning under the reading and listening conditions. On the other hand, Vidal's (2011) research on incidental learning of single words in academic lectures showed that listening resulted in significantly smaller learning gains than reading. This indicates that in the case of collocations in academic lectures, listening may lead to smaller learning gains than reading. Given the contrasting hypotheses, further research is warranted. Meanwhile, incidental learning research using graded readers (Webb & Chang, 2020) and storytelling videos for children as L2 input (Teng, 2019) revealed that reading while listening resulted in greater collocational learning than reading and listening, and viewing with captions led to greater learning gains than viewing without captions. If academic lectures were used as L2 input, the same patterns might be found. No empirical studies have been conducted to test this hypothesis. In fact, there is a lack of research investigating whether collocations in academic lectures can be learned through different input modes and comparing the learning gains across these modes.

The present study was carried out to address these gaps. Its primary aim was to compare the learning of collocations in academic lectures across five input modes: reading, listening, reading while listening, viewing, and viewing with captions. Its secondary aim was to explore the effect of frequency of occurrence, strength of association, type of vocabulary, type of verbal elaboration, nonverbal elaboration, comprehension, and prior knowledge of general vocabulary on the learning of collocations through these modes. The findings of this study would provide a thorough assessment of the potential of academic lectures for incidental learning of collocations. They also shed light on the relative value of written input, audio input, and visual input on the

learning of collocations in academic lectures as well as the effects of various factors on vocabulary learning through these input modes. Together these findings provide EAP teachers and learners with useful implications to make better use of academic lectures as a resource for incidental L2 vocabulary learning.

## Background

### Incidental learning of collocations

This study defines collocations as two-word combinations whose frequency of co-occurrence is statistically higher than by chance irrespective of their semantic relationship or level of compositionality ( Gablasova, Brezina, & McEnery, 2017). There are two reasons for choosing this frequency-based approach. First, unlike the phraseological approach which relies on human judgment of semantic relationship and level of compositionality of collocations, the frequency-based approach provides researchers with a quick and objective way to identify collocations (Boers & Webb, 2018). Second, frequency of occurrence is an important factor when selecting words for learning. It helps to draw learners' attention to the lexical items that they are likely to encounter often in their language use, and therefore, ensures effective selection (Nation, 2016). Knowledge of collocations is crucial for high levels of L2 proficiency (Nation, 2013), but many learners have insufficient collocational knowledge (Nguyen & Webb, 2017). Although collocations can be learned through explicit learning activities ( Boers, Demecheleer, & Eyckmans, 2004), it might be challenging for teachers to teach all collocations in class. Therefore, apart from explicit teaching, it is equally important to identify sources for learners to learn collocations incidentally.

Incidental vocabulary learning is a by-product of meaning-focused activities such as reading novels or watching television programs (Ellis, 1999). This means that when learners complete activities, their focus is on understanding the meaning of the input rather than deliberately learning a set of lexical items. Appendix S1 presents detailed summaries of previous research on incidental learning of collocations. The findings of studies examining incidental learning of collocations from reading written input are inconclusive. Szudarski (2012) and Szudarski and Carter (2016) found no significant learning of collocations through reading stories, but Pellicer-Sánchez (2017) did. Other studies also reported significant learning of collocations through reading graded readers (Webb & Chang, 2020). The findings of studies with aural input and audiovisual input, however, are more consistent. Collocations could be learned through reading while listening to graded readers (Webb, Newton, & Chang, 2013), listening to songs (Pavia et al., 2019), listening to teacher talk (Jin & Webb, 2020), viewing a short video clip without captions (Puimège & Peters, 2019), viewing a full-length television program without captions (Puimège & Peters, 2020), and viewing an academic lecture without captions (Dang et al., 2021). Only Teng (2019) and Webb and Chang (2020) have explicitly compared the learning of collocations through different input modes. Teng found that viewing storytelling videos for children with full captions always led to greater learning gains than viewing with keyword captions and viewing without captions. Webb and Chang found that reading while listening led to greater learning gains than the other input modes, and no significant difference was seen between reading and listening.

It should be noted that research on incidental learning of collocations has mainly focused on non-academic input (e.g., graded readers, television programs). Yet academic lectures are also potential sources for incidental learning of collocations for several reasons. First, corpus-driven

studies have revealed that academic lectures contain many frequently occurring sequences of words ( Biber, Conrad, & Cortes, 2004; Coxhead, Dang, & Mukai, 2013; Simpson-Vlach & Ellis, 2010). Moreover, the close relationship between topic-related vocabulary and academic context also means that academic input such as academic lectures may draw attention from learners (Coxhead, 2018). Additionally, in recent years, many universities have made their courses available as open access, which means that a large number of authentic academic lectures are now freely available to L2 learners worldwide. As repetition of topic-related vocabulary, learners' attention to input, and availability of resources are important elements in the selection of materials for meaning-focused language activities (Godfroid et al., 2018; Webb, 2015), academic lectures should be a useful source of L2 input for learning collocations.

Research on incidental learning using academic lectures as the L2 input has mainly focused on single words and only examined the learning under the listening (Vidal, 2003, 2011) and viewing without captions conditions (Dang et al., 2021; Smidt & Hegelheimer, 2004; Yang & Sun, 2013). These studies indicated that vocabulary learning gains may vary according to the input mode (see Appendix S2 for detailed summaries of these studies). Only Dang et al. (2021) have investigated the learning of collocations in academic lectures. They found that viewing an academic lecture could lead to the learning of 4.68 out of 19 collocations at the form recognition level. Dang et al. (2021) only examined academic lectures through viewing without captions. Recently thanks to the growing number of open access courses delivered by well-known universities, academic lectures are freely available in different formats (transcripts, audios, videos, and videos with captions). To provide a better assessment of the potential of academic lectures for incidental learning of collocations, it is important to investigate the learning of collocations in academic

lectures through different input modes (reading, listening, reading while listening, viewing, and viewing with captions) and compare the learning gains across these modes.

Comparing incidental learning of collocations through reading and listening to academic lectures is useful. On the one hand, Webb and Chang's (2020) findings on the learning of collocations in graded readers indicate that there might be no significant differences in the learning gains of collocations in academic lectures under the listening only and reading only conditions. On the other hand, Vidal's (2011) findings on the learning of single words in academic lectures suggest that listening would probably lead to smaller learning gains of collocations in academic lectures than that reading. Given these contrasting hypotheses, empirical studies comparing the learning of collocations in academic lectures under the listening only and reading only conditions would shed light on this issue.

It is important to investigate the learning of collocations through reading while listening and viewing an academic lecture with and without captions. Webb and Chang (2020) found that reading while listening to graded readers better facilitated the learning of collocations than reading only and listening only. Teng (2019) found that viewing storytelling videos for children with captions led to greater learning gains of collocations than viewing without captions. Therefore, if academic lectures are used as L2 input, reading while listening would probably lead to greater collocational learning gains than reading only and listening only, and viewing with captions may result in greater learning gains than viewing without captions. As no empirical studies have been conducted to test these hypotheses, further research is warranted.

**The multimedia learning theory and other influential SLA learning theories**

Recently the cognitive theory of multimedia learning has drawn a growing attention from SLA research (Mayer, 2014). This theory was built on Paivio's (1986) Dual-Coding Theory, which suggests that human process information through two different channels, which have limited processing capacities. There are two different approaches toward identifying these two channels. Based on the format of the stimulus, the presentation-mode approach categorizes these channels into verbal channel (e.g., aural and written forms of a word) and nonverbal channel (e.g. pictures, graphs). Meanwhile, the sensory-modality approach is based on the stimulus and classifies the two channels into auditory channel (processing the information through the ears) and visual channel (processing the information through the eyes). According to Mayer (2014), the multimedia learning theory proposes that the learning outcome can be enhanced if students construct information from both channels rather one channel, because this would take advantage of students' full capacity for processing information. However, Mayer (2014) also calls for further research to determine whether the presentation-mode approach or the sensory-modality approach better distinguishes the two channels.

Two important principles of the multimedia learning theory are the split-attention principle and the redundancy principle. The split-attention principle suggests that multimedia learning materials should avoid requiring students to split their attention between multiple sources of information otherwise it will create extraneous cognitive load, which will negatively affect the learning (Ayres & Sweller, 2014). Related to the split-attention principle is the redundancy principle which states that redundant materials (i.e., materials that duplicate the same information in multiple forms or are unnecessarily elaborated) do not facilitate but interfere with learning (Kalyuga & Sweller, 2014). This is supported by the cognitive load theory (Sweller et

al., 1998), which states that human working memory is limited. When students are presented with several sources of unnecessary information spontaneously (e.g., pictures and words together with the aural and written forms of the words), they need to coordinate the redundant information from these sources, which would increase extraneous cognitive load for their working memory and inhibit learning (Kalyuga & Sweller, 2014).

There is a solid body of evidence in the context of learning of science in L1 supporting the multimedia learning theory (i.e., better learning outcomes are achieved if the information is processed in both channels rather than one channel) and its principles (i.e., duplicating the same information in multiple forms hinders learning) (Mayer, 2014). Studies examining the role of multimodal input on L2 learning also support this theory by showing that multimodal input (graded readers, television programs) led to learning (e.g., Puimège & Peters, 2019, 2020; Peters & Webb, 2018). However, at the same time these studies provide evidence contrasting the split-attention principle and the redundancy principle. They found that duplicating the same information in multiple forms – reading while listening to graded readers/stories (i.e., aural and written forms of the words were both presented) and viewing television programs with captions and subtitles (i.e., aural, visual, and written forms of the words were all presented) – did not hinder comprehension nor learning, including vocabulary learning (e.g., Pellicer-Sánchez et al. 2020; Peters, 2019; Teng, 2019; Webb & Chang, 2020).

Based on these findings, SLA researchers have suggested that perhaps L1 students and L2 students process multimedia input differently. The combination of aural input with written input (reading while listening) and visual input (viewing with captions and viewing with subtitles) allows L2 learners to decode and segment the speech stream and support their listening comprehension, which would increase the chance of noticing unfamiliar vocabulary in the input

(Montero Perez, 2020). It is important to note that studies supporting the split-attention principle and the redundancy principle used academic materials with a high level of complexity, mainly from science and mathematics, to teach subject-specific content. In contrast, studies that did not provide evidence supporting these principles used non-academic input with a lower level of complexity (e.g., graded readers, television programs, storytelling videos for children). The higher level of complexity of academic input may make it more challenging for L2 learners to integrate information from different channels when they are exposed to academic input than when they are exposed to non-academic input. Especially, in the case of academic lectures, the challenge may be even greater, because learners have to process a large amount of complex and abstract information spontaneously over a long period of time (Biber, 2006; Flowerdew, 1994), which might create a greater burden on their cognitive system and make it difficult for them to integrate information from different channels. Therefore, it would be useful to investigate whether the split-attention principle and the redundancy principle hold true when L2 learners process multimedia input with high complexity such as academic lectures.

Apart for the multimedia learning theory, several other theories also have a great influence on SLA research and may provide further insights into L2 vocabulary learning through multimodal input. According to Krashen's (1985) input hypothesis, learning happens if students are exposed to a large amount of comprehensible input which is slightly above their current level of competence (Feng, 2021). As a large number of academic lectures from many universities are available as open access, these lectures are potential comprehensible input for EAP learners as well as L2 learners studying at English medium universities. Meanwhile, Schmidt's (1990) noticing hypothesis suggests that learners only learn particular linguistic features in comprehensible input if they notice them in the input. Additionally, Craik and Lockhart's (1972)

depth of processing hypothesis proposes that the deeper learners process the information in the input, the more likely learning happens. However, VanPattern's (2007) input processing theory suggests that L2 learners have limited processing capacity; therefore, they can only process a certain amount of information and attend to certain linguistic patterns at a given period of time. As open access academic lectures can be presented to learners through different input modes (reading, listening, reading while listening, viewing, and viewing with captions), it would be useful to explore which modes impose the lightest cognitive load on L2 learners and create the best condition for them to notice and learn unfamiliar collocations in academic lectures.

## Factors affecting incidental learning of collocations

Various factors may affect the chance of lexical items being noticed and learned through comprehensible input (Peters, 2020; Webb, 2020). Examining the relative contribution of these factors on incidental learning of collocations in academic lectures across different input modes would help teachers to be better aware of the factors that facilitate or hinder the learning of collocations in academic lectures and make better use of this kind of input. No attempts have been made to address this gap. However, findings of previous research on incidental learning of single words in academic lectures and incidental learning of collocations on non-academic input suggest that frequency of occurrences, strength of associations, elaboration, type of vocabulary, comprehension, and prior vocabulary knowledge may potentially affect incidental learning of collocations when an academic lecture is used as L2 input.

### *Frequency of occurrence*

According to the input hypothesis (Krashen, 1985), for incidental vocabulary learning to happen through meaning-focused input, learners need to have repeated exposure to the target vocabulary (Feng, 2021; Web & Nation, 2017). Therefore, frequency of occurrences of vocabulary in L2

input has received a great deal of attention from incidental vocabulary studies. Research has found that frequency of occurrence positively affects the learning of single words through reading (e.g., Vidal, 2011), listening (e.g., Pavia et al., 2019; van Zeeland & Schmitt, 2013; Vidal, 2003, 2011), and viewing (e.g., Peters, 2019; Peters & Webb, 2018; Yang & Sun, 2013). In fact, Uchihara, Webb, and Yanagisawa's (2019) meta-analysis revealed that frequency significantly contributed to the incidental learning of single words. Compared to the considerable number of studies with single words, studies examining the effect of frequency on incidental learning of collocations are limited in number, and their findings have been inconsistent. Research has revealed significant, positive correlations between frequency and learning gains of collocations through reading (Webb & Chang, 2020), listening to (Webb & Chang, 2020), and reading while listening to graded readers (Webb & Chang, 2020; Webb et al., 2013). However, studies also found no significant correlations through reading stories (Pellicer-Sánchez, 2017), listening to teacher talk (Jin & Webb, 2020), and viewing academic lectures (Dang et al., 2021). As the frequency of occurrence of multiword sequences varies according to modality and kind of discourse (Biber, 1988; 2006), the conflicting results suggest that the effect of frequency on incidental learning of collocations might be different depending on the kind and mode of input. As only Dang et al. (2021) has investigated academic lectures as the source of input, further research exploring the impact of frequency on the learning of collocations in academic lectures through different input modes is warranted. Such research would provide further insights into the amount of input needed for incidental learning of collocations in academic lectures to happen and extend our current knowledge of the effect of frequency on incidental vocabulary learning.

### *Strength of association*

Strength of association refers to the likelihood of the two words that make up collocations co-occurring with each other (Gablasova et al., 2017). The pace of learning a collocation is likely to be affected by the degree of substitutability of its constituents (Boers et al., 2014). Research on L2 learner knowledge of collocations has consistently found that collocations consisting of high-frequency components tend to be known better than those consisting of low-frequency components (e.g., Nguyen & Webb, 2017). Therefore, the strength of associations may have a significant impact on noticing, which would then affect the chance of a collocation being learned. Only Puimège and Peters (2019, 2020) have investigated the effect of strength of association on incidental learning of collocations. Their findings were fairly mixed. Both studies used television programs as input and used mutual information (MI) scores to represent strength of association. Puimège and Peters (2019) found that the MI scores negatively correlated with the learning gains at the form recall level. However, Puimège and Peters (2020) found that MI scores did not significantly correlate with learning gains at the form recall level, but positively and significantly correlated with learning gains at the meaning recall level. As very few studies have examined the effect of the strength of association on incidental learning of collocations and no studies have explored the effect of this factor on the learning of collocations in academic lectures through different input modes, further research is needed. Such research would provide teachers with implications on the learning burden faced by learners when learning collocations in academic lectures.

### *Verbal elaboration and non-verbal elaboration*

According to the noticing hypothesis (Schmidt, 1990) and the depth of processing hypothesis (Craik & Lockhart, 1972), the deeper learners process the information, the more likely they

attend to and learn the items. Academic lectures are the speech events in which lecturers inform, explain, evaluate, and critique key information related to the course content to their students[1] (Dang et al., 2017; Lynch, 2011). Therefore, elaboration provided by lecturers may provide opportunities for learners to process the information deeply, which would increase chances for collocations in the lectures to be noticed and learned. Elaboration includes verbal elaboration and non-verbal elaboration. According to Chaudron (1982), verbal elaboration can be divided into explicit elaboration and implicit elaboration. Explicit elaboration means lexical terms are explained in the form of definitions, descriptions, naming, and questioning statements. Implicit elaboration means lexical terms are explained through examples, paraphrase, or synonymy. Several studies have found that the type of verbal elaboration significantly affected the learning of single words through listening to (Vidal, 2003, 2011) and viewing academic lectures (Yang & Sun, 2013). That is, explicit verbal elaboration led to the largest vocabulary learning gains, followed by implicit verbal elaboration; words without any verbal elaboration had the least vocabulary learning gains. Nonverbal elaboration means that lexical terms are explained through gestures, symbols, or images (Yang & Sun, 2013). Yang and Sun (2013) investigated the impact of this factor on the learning of single words when academic lectures were used as L2 input and found that words with nonverbal elaboration were more likely to be learned than words without nonverbal elaboration. No studies have examined the effect of elaboration on incidental learning of collocations in general and incidental learning of collocations in academic lectures in particular. Such research is important because it would provide useful implications on the kinds of elaboration that content lecturers should use to better facilitate L2 learners' incidental learning of collocations.

### *Type of vocabulary*

The type of vocabulary might contribute to incidental learning of collocations in academic lectures because it may affect the levels of attention and processing new vocabulary (Peters & Webb, 2018). Research has shown that the type of vocabulary significantly affected incidental learning of single words through listening to (Vidal, 2003, 2011) and viewing academic lectures (Yang & Sun, 2013). Technical words were the most likely to be learned, followed by low-frequency words, and then academic words. Whether the type of vocabulary influences incidental learning of collocations in academic lectures and how its effect varies across different input modes remain to be answered. Moreover, previous research investigating the effect of the type of vocabulary on incidental learning of single words (Vidal, 2003, 2011; Yang & Sun, 2013) considered specialized vocabulary (academic words and technical words) being different from low-frequency vocabulary. Recently researchers (Dang, 2020; Dang et al., 2017; Gardner & Davies, 2014; Nation, 2016) have pointed out that specialized vocabulary can also be high, mid, and low-frequency words. Therefore, there is a need for further research exploring the effect of the type of vocabulary on incidental learning of collocations across different input modes and such research should adopt the current approach toward defining specialized vocabulary. Findings of such research would shed light on the kinds of collocations in academic lectures that are more likely and less likely to be learned incidentally.

### *Comprehension*

Comprehensible input is important for incidental language learning (Krashen, 1985). It creates conditions for learners to notice the forms of unfamiliar lexical items in the input and infer their meanings from context. The significant role of comprehension in language learning has led to the hypothesis that the better learner's comprehension of the input, the greater the vocabulary

learning gains from the input. However, while a great deal of research has examined the effect of the percentage of known words in texts (e.g., Schmitt et al., 2011; van Zeeland & Schmitt, 2013) and different kinds of L2 input (e.g., Montero Perez et al., 2014; Smidt & Hegelheimer, 2004; Teng, 2019) on comprehension, empirical studies exploring the effect of comprehension on incidental vocabulary learning gains is surprisingly rare. Only Pulido (2007) has attempted to address this gap. Focusing on learners of Spanish as an L2, she found that comprehension significantly contributed to vocabulary learning gains at the form recognition, meaning recognition, and meaning recall levels. This finding supports the hypothesis that vocabulary learning gains increase according to the degree of comprehension. However, Pulido (2007) examined the learning of single words through reading narrative passages. It is unclear whether the same result would be found if collocations and different kinds of input and modes of input are examined.

***Prior vocabulary knowledge***

Research has consistently shown that prior vocabulary knowledge positively affects the learning of single words through reading (e.g., Horst, Cobb, & Meara, 1998), reading while listening to graded readers (Webb & Chang, 2015a, 2015b), and viewing television programs (e.g., Peters, 2019; Peters & Webb 2018), and the learning of collocations through viewing televisions (Puimège & Peters, 2019, 2020). Fewer studies have focused on academic input. Vidal (2003, 2011) found that language proficiency positively affected the learning of single words through listening to academic lectures and reading academic texts, but she did not specifically examine the impact of prior vocabulary knowledge on learning. Meanwhile, Dang et al. (2021) found that prior knowledge of general vocabulary did not make any significant contribution to the learning of single words nor collocations through viewing an academic lecture. However, Dang et al.

(2021) only focused on viewing without captions. The impact of prior vocabulary knowledge on incidental learning of collocations may vary depending on the input mode because certain input modes may facilitate incidental learning of collocations better than the others (Webb & Chang, 2020). Yet, no studies have investigated the effect of prior knowledge of general vocabulary on incidental learning of collocations in academic lectures across different input modes. Addressing this gap would reveal how learners' existing vocabulary knowledge may influence collocational learning when academic lectures are presented in different formats. It would then provide teachers with useful implications on how to make better use of academic lectures to suit learners of different vocabulary knowledge.

## Research questions

Despite the potential of academic lectures as resources for incidental learning of collocations, no studies have compared the learning of collocations in academic lectures through different input modes. Nor has any research explored the effects of frequency of occurrence, strength of association, verbal elaboration, nonverbal elaboration, type of vocabulary, comprehension, and prior knowledge of general vocabulary on the learning through these modes. This study would address these gaps by investigating three questions:

1. To what extent do L2 learners incidentally learn collocations in an academic lecture through reading, listening, reading while listening, viewing, and viewing with captions?
2. How do collocation learning gains compare across these input modes?
3. What is the relationship between the collocation learning gains and (a) frequency of occurrence, (b) strength of association, (c) type of verbal elaboration, (d) nonverbal

elaboration, (e) type of vocabulary, (f) comprehension, and (g) prior knowledge of general vocabulary?

## Methodology

The participants were 165 postgraduate students majoring in Technology and Engineering from six intact classes in an EAP course at a university in China. Each group of participants was randomly allocated to a control group and five experimental groups. Further information about the participants is present in Appendices S3 and S4. Each experimental group encountered the target collocations in the same academic lecture in one of the following input modes: (a) reading, (b) listening, (c) reading while listening, (d) viewing, and (e) viewing with captions. Webb, Sasao, and Balance's (2017) Updated Vocabulary Levels Test (UVLT) was administered two weeks before the treatment as part of the participants' program entry test. One-way between groups ANOVA showed no significant difference between the means scores of the six groups on the UVLT: $F$ (5,159) = 1.5, $p$= .19, $\eta^2$= .05, and the English test of the Master Entrance Examination: $F$ (5,136) = 1.77, $p$= .12, $\eta^2$= .02. This result indicates that the six groups had similar knowledge of the most frequent 3000 words and similar language proficiency.

The materials were the video, audio, and transcript of an introductory undergraduate lecture in an open access course in Algorithms delivered by Massachusetts Institute of Technology (see Appendix S5 for the transcript of the lecture and Appendix S6 for the detailed steps of the material selection). Nineteen collocations[2] that occurred in the lecture were selected as the target items (see Appendix S7 for these collocations and Appendix S8 for the steps of selecting them and identifying their frequency, strength of association, vocabulary types, and elaboration).

A multiple-choice test was created to measure knowledge of the target collocations (Appendix S9). Participants from all groups listened to the aural clues and read the written clues of all options so that the test results would not be biased towards a particular input mode. To complete the test, the participants have to select among the three options (the target collocations and two distracters) the two words that were likely to appear together in English. If they did not know, they could choose an '*I don't know'* option (see Figure 1).

*Figure 1. Example of a test item*

The participants saw on paper:

*Q1 a integral work b integral part c integral lab d I don't know*

At the same time they heard on the recording:

| | | | |
|---|---|---|---|
| *Question 1* | *[pause for 1 second]* | | |
| *a* | *[pause for 1 second]* | *integral work* | *[pause for 2 seconds]* |
| *b* | *[pause for 1 second]* | *integral part* | *[pause for 2 seconds]* |
| *c* | *[pause for 1 second]* | *integral lab* | *[pause for 5 seconds]* |

Examples were provided and instructions were given in Chinese so that the participants understood what to do in the test. The test was used as the pretest, immediate posttest, and delayed posttest, but the order of the items in these tests were different to minimize testing effects. One collocation (*computer science*) known by at least 80% of the students in a pilot study was also included in the test to encourage the participants to complete the test, but it was removed from the analysis. The internal consistency of the test was good (Cronbach's $\alpha = .74$).

A comprehension test of 10 True/False items was delivered at the end of the treatment to measure the participants' understanding of the lecture (see Appendix S10 for the test and Appendix S11 for the steps taken when designing the test).

The experiment was conducted over three weeks. In the first week, the participants were told that the purpose of the study was to investigate the effect of input modes on academic listening comprehension and all groups completed the pretest. In the second week, the experimental group completed the treatment in the assigned input modes. Right after the treatment, all experimental groups completed the comprehension test and the immediate posttest. The control group did not complete a treatment and did the immediate posttest only. In the last week, all groups completed the delayed posttest. The real purpose of the study was revealed to the participants after they had handed in their delayed posttest.

Data were scored dichotomously. Correct answers were given 1 and incorrect answers were given 0. Statistical analyses were carried out with R statistical platform. A single linear mixed-effects model was constructed to find the answer to Research questions 1 and 2, and a series of general linear mixed effects models were performed to find the answer to Research question 3 (see Appendix S12 for the steps of the analysis).

## Results

### Learning gains and retention through different input modes

The mean scores of all groups increased from the pretest to the immediate posttest and from the pretest to the delayed posttest (see Appendix S13). The analysis with the single linear mixed-effects model revealed that the fixed effects (mode, time, mode by time interaction, and vocabulary level) explained 36% of the variance in the scores (marginal $R^2$= .36) and the whole model (both the fixed effects and random effects) explained 63% of the variance (conditional $R^2$ = .63). Appendix S14 shows that there was a significant main effect of time on the scores ($p$ < .001), indicating that the overall mean scores of the participants significantly increased from

the pre-test to the post-tests. The mode of input did not have a significant main effect on the scores in most cases ($p > .05$), except for viewing ($p = .002$). We found significant mode by time interaction at the $p < .05$ level in the cases of reading, viewing, and viewing with captions, but not in the case of listening and reading while listening for the immediate post-test. This means that the reading, viewing, and viewing with captions had significantly higher mean scores on the immediate post-test than the control group while listening and reading while listening did not. We found no significant group by time interaction at the $p < .05$ level for the delayed post-test, suggesting that none of the experimental groups had significantly higher mean scores on the delayed post-test than the control group.

To further examine the interaction between mode and time, we ran a series of pairwise comparisons test using the emmeans package in R with Bonferroni adjustments for multiple comparisons. When the scores at three testing times of each group were compared, in most cases, the immediate posttest scores and the delayed posttest scores were significantly higher than the pretest scores (all $p<.05$), but there was no significant difference between the immediate posttest and delayed posttest scores (all $p> .05$). The only exception is the reading while listening group whose immediate posttest scores were significantly higher than the pretest scores ($p < .0001$) and the delayed posttest scores ($p < .0001$), but no significant differences were found between the pretest scores and the delayed posttest scores ($p= 1.0$). These findings suggest that learning happened in all conditions.

The pairwise comparison also revealed no significant differences between the pre-test scores of the six groups (all $p>.05$), indicating that the control group and the five experimental groups had similar knowledge of the target collocations before the treatment. When the immediate post-test scores of the six groups were compared, the scores of the control group were not significantly

different from those of the listening *(p*= 1.0) and reading while listening groups (*p*= *.07*), but were significantly lower than those of the reading (*p* = .046), viewing (*p*<.0001*)*, and viewing with captions groups (*p*= *.02*). Moreover, no significant differences were found in the immediate post-test scores of the reading, viewing and viewing with captions groups (all *p*>.05). This suggested that listening and reading while listening did not lead to learning of collocations. However, reading, viewing, and viewing with captions did, and these three conditions had equal potential for incidental learning of collocations from academic lectures. For the delayed post-test scores, the control group had significantly lower scores than the viewing group (*p*=.006), but no significant differences were found between the delayed post-test scores of the control group and the other experimental groups. When we relevelled the model using the reading, listening, reading while listening, viewing, and viewing with captions groups in turns as the reference level for mode, the results were consistent with those found when the control group was used as the reference level.

**Effects of the seven variables on incidental learning of collocations**

For ecological validity, we did not manipulate the lecture nor the target collocations. Therefore, the findings presented in this section should not be generalized, but should be treated as the starting point for further investigation. Analysis of the data of all five groups combined with the general linear mixed-effects models revealed that the whole model explained 5% of the variance in the scores (conditional $R^2$ = .05) while the fixed effects (frequency of occurrence, strengths of association, type of vocabulary, type of verbal elaboration, nonverbal elaboration, comprehension scores, and prior knowledge of general vocabulary) explained 3% of the variance (marginal $R^2$ = .03).

Appendix S15 presents the results of the analysis. Nonverbal elaboration significantly affected

the immediate posttest scores of the five experimental groups combined ($p < .001$). The log (odds) of a correct response was higher for a collocation with nonverbal elaboration than a collocation without nonverbal elaboration. This indicates that nonverbal elaboration had a significant effect on the learning of all five experimental groups combined. However, frequency, MI score, the type of vocabulary, the type of verbal elaboration, comprehension, and prior vocabulary knowledge did not significantly affect immediate posttest scores of the five experimental groups combined (all $p > .05$). When we relevelled the model to directly compare different kinds of vocabulary (administrative, academic, and technical) and verbal elaboration (implicit and explicit), similar results were found. This suggests that frequency of occurrence, strength of association, type of vocabulary, type of verbal elaboration, comprehension, and prior knowledge of general vocabulary made no significant contribution to the learning gains of the five experimental groups combined, but nonverbal elaboration did.

Analysis of each experimental group showed that none of the examined factors significantly affected the learning gains of the reading, reading while listening, and viewing groups (all $p > .05$); however, several factors significantly influenced the learning of the listening and viewing with captions groups. Appendix 16 presents the results of the listening group. Nonverbal elaboration significantly contributed to the learning gains. The log (odds) for a collocation with nonverbal elaboration to be learned was higher than that for a collocation without nonverbal elaboration. Frequency, MI scores, the type of verbal elaboration, comprehension, and prior vocabulary knowledge did not significantly affect the learning gains (all $p > 0.05$). When the model was relevelled to directly compare explicit and implicit elaboration, the same results were found. Appendix S16 also shows that compared to non-specialized collocations, the log (odds) of a correct response for administrative collocations, academic collocations, and technical

collocations was not significantly different (all $p > 0.05$). Yet when the model was leveled to directly compare administrative, academic, and technical collocations, the log (odds) for academic collocations to be learned was significantly higher than that for technical collocations ($p$ =0.03). In brief, the findings suggest that nonverbal elaboration and the type of vocabulary had a significant effect on the collocation learning of the listening group, but frequency, strength of association, type of verbal elaboration, comprehension, and prior vocabulary knowledge did not.

Appendix S17 presents the results of the viewing with captions group. Frequency, MI, type of vocabulary, nonverbal elaboration, comprehension, and prior vocabulary knowledge did not significantly affect the learning gains (all $p$>.05).When the models were relevelled with administrative vocabulary, academic vocabulary, and technical vocabulary taking turns as the reference level for the type of vocabulary, similar results were found. In terms of verbal elaboration, as shown in Appendix 17, the chances for collocations with implicit elaboration and those with explicit elaboration to be learned were not significantly different from those without verbal elaboration. However, when the model was relevelled to directly compare explicit and implicit elaboration, the log (odds) for a collocation with explicit elaboration to be learned was higher than that for implicit elaboration ($p$= .02). In sum, the type of verbal elaboration significantly affected the collocation learning of the viewing with captions group, but frequency, strength of association, type of vocabulary, nonverbal elaboration, comprehension, and prior vocabulary knowledge did not.

## Discussion

The present study extends research on incidental vocabulary learning in many ways. It is among the very few studies that have investigated incidental learning of collocations. In fact, it is the first to (a) investigate the learning of collocations in academic lectures through different input modes, (b) compare the learning across these modes, and (c) explore the effect of frequency of occurrences, strength of association, type of vocabulary, elaboration, comprehension, and prior knowledge of general vocabulary on the learning of collocations across these modes.

### The role of aural input, written input, and visual input in incidental learning of collocations in academic lectures

In answer to the first two research questions, the results showed that collocational learning happened for all five experimental groups as well as the control group. This finding indicated that the learning gains were not attributed solely to the treatments, but might also be the results of the exposure to the target collocations on the pretest. Testing effects are common in incidental vocabulary learning research (e.g., Jin & Webb, 2020; Webb et al., 2013). That is why it is essential to include a control group to ensure that the gains are attributed solely to the treatments. The present study found that the listening and reading while listening groups did not have significantly greater collocational learning gains than the control group whereas the reading, viewing, and viewing with captions groups did. These results suggested that listening and reading while listening to an academic lecture did not result in a significant improvement in the learning gains of collocations at the form recognition level, but reading, viewing, and viewing with captions did.

This study found that on average, the number of collocations learned by the reading, viewing, and viewing with captions groups in turn was 4.88 out of 19 collocations (25.68%), 4.68

collocations (24.63%), and 4.42 collocations (23.26%), and that the learning gains across these three input modes were not significantly different. These findings are in line with Dang et al.' (2021) finding that collocations could be learned through viewing an academic lecture. However, expanding on Dang et al.' study, this study showed that this trend also holds true in the case of reading and viewing with captions, but not in the case of listening and reading while listening. To some extent, this finding contrasts with the findings of previous studies on incidental learning of collocations. Teng (2019) found that viewing storytelling videos for children with captions resulted in greater gains than viewing without captions. Meanwhile, Webb and Chang (2020) found that reading while listening to graded readers led to greater learning gains than reading only and listening only, and there were no difference in the gains through reading only and listening only. One possible reason for the conflicting findings is that Teng (2019) and Webb and Chang (2020) examined non-academic input while the present study investigated academic input. The findings of the present study suggest that aural input, written input, and visual input may have distinctive contributions to the learning of collocations in academic lectures.

To begin with, aural input alone does not seem to significantly contribute to the learning of collocations in academic lectures. This study found no significant difference in the learning gains of the listening group and the control group, which contrasts with Vidal's (2003, 2011) finding that listening to academic lectures led to learning of single words. The conflicting results might be because Vidal (2003, 2011) used 15-minute lectures which had been modified from authentic sources and recorded by a lecturer who 'rephrased some of the information as if she were interacting with the listeners' (Vidal, 2003, p.62) rather than a lecture that was recorded in an actual class. In contrast, the present study used an unmodified academic lecture of 50 minutes. When listening to authentic academic lectures, students have to process a large amount of dense

and abstract information spontaneously over a long period of time (Biber, 2006; Flowerdew, 1994). Moreover, apart from the challenges caused by academic listening, L2 learners also need to deal with the challenges caused by the features of spoken English (e.g., connected speech, irregular pauses, false starts, hesitations, stress, and intonation) in the process of decoding the meaning of L2 input (Field, 2011; Flowerdew, 1994; Goh, 2000). These features may be even more problematic for L2 learners coming from an educational context with more focus on written input than aural input (Flowerdew, 1994) and those from EFL contexts in which the amount of aural L2 input is smaller than that in ESL contexts (Webb & Nation, 2017). According to the noticing hypothesis (Schmidt, 1990), unfamiliar linguistic forms are only learned in comprehensible input if they are noticed. Meanwhile, the input processing theory (VanPatten, 2007) suggests that learners can only attend to unfamiliar linguistic forms if they have sufficient attentional resources. When the participants listened to the academic lecture, the demand caused by academic listening and L2 listening may create a great burden on their cognitive system. As a result, they might be busy trying to understand the message in the lecture rather than attending to the linguistic features in the input. This may explain why collocations were not learned through listening to an unmodified academic lecture.

The current study also suggests that written input alone is likely to make a significant contribution to the learning of collocations in academic lectures. As shown in this study, the reading group had significantly higher collocation learning gains than the control group. This finding is consistent with earlier findings that collocations could be learned through reading stories (Pellicer-Sánchez, 2017) and graded readers (Webb & Chang, 2020). The greater role of written input than aural input in the learning of collocations in academic lectures might be because unlike aural input, written input offers learners the opportunities to pause and refer back

to parts of the lectures which they do not understand and adapt the perceptual process to match their needs (Flowerdew, 1994). As a result, the cognitive load of processing written input might not be as great as that of aural input, which may make unfamiliar vocabulary in written input more salient to learners than in aural input.

This study also suggests that not all kinds of input when combined with aural input can support the learning of collocations in academic lectures. Visual input seems to have a positive contribution to learning through aural input. As shown in this study, aural input accompanied with visual input – viewing (audiovisual input) and viewing with captions (audiovisual plus written input) – resulted in significant learning gains. In contrast, aural input accompanied with written input does not appear to make any significant difference in collocation learning through aural input. In this study, neither aural input alone (listening) nor aural input plus written input (reading while listening) led to significant learning gains. Moreover, although both audiovisual input (viewing) and audiovisual plus written input (viewing with captions) resulted in significant learning gains, the gains from the latter were not significantly higher than those from the former.

The relative value of visual input and written input in supporting aural input found in the present study helps to expand on the multimedia learning theory in several ways. First, it suggests that in the case of academic lectures, the presentation-mode approach is probably a better way to classify the two channels for processing information from multimedia input than the sensory-modality approach. According to the multimedia learning theory, better learning outcomes are likely to be achieved if information from multimedia learning materials is processed in both channels rather than in one channel. The presentation-mode approach classifies written input and aural input as being processed in the same channel (verbal) and visual input in another channel (nonverbal). Meanwhile, the sensory-modality approach categorizes aural input as being

processed in one channel (auditory) whereas written input and visual input as being processed in the other (visual). This study found that combining visual input and aural input significantly contributed to the learning, but combing written input and aural input did not, which supports the presentation-mode approach.

Second, this study provides evidence supporting the split-attention principle and the redundancy principle. According to these principles, processing information from multimedia learning materials through only one channel would create extraneous cognitive load for learners' working memory and hinder learning; however, processing information from these materials through both channels would lighten the cognitive load and facilitate learning. This study found that combining written input and aural input, which were processed in the verbal channel, did not lead to significant learning gains. In contrast, combining visual input and aural input, which were processed in two separate channels (verbal and nonverbal), did.

By supporting the split-attention principle and the redundancy principle, the present study contrasted previous SLA studies which did not support these principles (e.g., Pellicer-Sánchez et al. 2020; Peters, 2019; Webb & Chang, 2020). However, it should be noted that previous SLA studies examined L2 vocabulary learning through non-academic input (graded readers and television programs) while the present study used an unmodified academic lecture on Algorithms. Academic materials are likely to have a higher level of complexity, and thus, may be more challenging for students to integrate information from different input modes than non-academic texts. As previously mentioned, when listening to the academic lecture (aural input only), the participants had to process a large amount of complex and abstract information spontaneously as well as decoding the meaning of the input in another language. This would create a great burden on their working memory. When written input was used to support aural

input, the information from both input modes were likely to be processed in the same channel, and the participants might have to split their attention between aural input and written input while still had to catch up with the speed of the lecture. This may then increase extraneous cognitive load for their working memory. Consequently, the participants may not be able to make the best use of written input to support vocabulary learning. In contrast, when visual input was used to support aural input, information from these input modes was likely to be processed in separate channels. As a result, the participants' capacity for processing information would be used in full, which would enhance their attentions and the learning of the collocations in the lecture.

The present study found that reading, viewing, and viewing an academic lecture with captions could lead to the learning of 4.88 (25.68%), 4.68 (24.63%), and 4.42 (23.26%) collocations at the recognition level, respectively. Readers may think that this amount was small and question the value of the academic lectures for incidental vocabulary learning. However, there are several reasons why the learning gains in the present study was meaningful. First, they are either larger than or comparable to those reported in previous studies on incidental learning of collocations at the form recognition level. The number of collocations learned through reading the script of an academic lecture (4.88 collocations) was larger than the number of collocations learned through reading graded readers (2.04 collocations; Webb & Chang, 2020) and stories (no significant learning gains; Szudarski, 2012; Szudarski & Carter, 2016, or 3.20 and 2.90 collocations; Pellicer-Sánchez, 2017). The number of collocations learned through viewing an academic lecture without captions (4.68 collocations) was also larger than the number of collocations learned through viewing a series of short videos for young children without captions (2.03 collocations; Teng, 2019). Additionally, the number of collocations learned through viewing an

academic lecture with captions was fairly comparable to those learned through viewing a series of storytelling videos with captions (4.89 and 6.88 collocations; Teng, 2019). Moreover, the learning gains of collocations found in the present study are also larger than the gains found in previous research on incidental learning of single words when academic lectures were used as L2 input: 16% (Vidal, 2003), 16% (Smidt & Hegelheimer, 2004), 22.69% and 15.48% (Vidal, 2011), 10.12% (Yang & Sun, 2013), and 8.14% (Dang et al., 2021). Another reason why the learning gains found in this study is meaningful is that this study only measured knowledge of the target collocations and only focused on knowledge at the form recognition level. L2 learners also learn other aspects of the target and non-target collocations when being exposed to meaning-focused input (Webb, 2020). The exposure to academic input such as lectures may also help learners to notice how known lexical items were used in specific contexts and expand their knowledge of the specialized meanings of these items (Coxhead, 2018). Therefore, it could be expected that the gains found in the present study is smaller than those actually occurred. Last but not least, the participants in this study were exposed to only one academic lecture and only read/view it once. Yet incidental learning is incremental in nature. Previous studies have revealed that exposure to multiple texts resulted in greater learning gains than exposure to a single text (Rodgers & Webb, 2020; Teng, 2019; Webb & Chang, 2015a, 2015b) and that repeated exposure to the same text led to greater gains than exposure to it once (Pavia et al., 2019). Therefore, if learners were exposed to a large number of lectures or read/view a lecture for multiple times, the gains may be greater.

**The relationship between the collocation learning gains and other factors**

Despite being exploratory in nature, the present study provides further insights into the effects of frequency of occurrence, strength of association, type of verbal elaboration, nonverbal

elaboration, type of vocabulary, comprehension, and prior knowledge of general vocabulary on collocational learning.

### *Frequency of occurrence*

This study did not find a significant effect of frequency of occurrence on the learning of collocations in academic lectures. This supports Dang et al.' (2021) finding that frequency did not significantly contribute to the learning of collocations through viewing an academic lecture. However, expanding on Dang et al.'s, this study showed that the same pattern was seen in reading, listening, reading while listening, and viewing with captions. The insignificant effect of frequency on vocabulary learning could be explained in two ways. First, the range in number of encounters (1-9) with the target collocations in the current study might not be sufficient for frequency to have a significant impact on learning. However, as this study used unmodified input, this finding may provide ecologically valid results on the learning of collocations in lectures. Second, learning from comprehensible input is affected by the level of attention (Schmidt, 1990) and the depth of processing (Craik & Lockhart, 1972). Other factors (e.g., elaboration, type of vocabulary) might play a more important role in the learning of collocations than frequency.

### *Strength of association*

This study found no significant effect of the strength of association on the learning of collocations in academic lectures. This contrasts with previous studies (Puimège & Peters, 2019, 2020) which reported a significant relationship between the strength of association and the learning gains through viewing television programs. There are three reasons for this difference. First, the impact of the strength of association might vary according to the kind of input. This study examined academic lectures while Puimège and Peters (2019, 2020) investigated television

programs. Second, this study purposefully selected collocations with MI scores of at least 3 as the target words whereas Puimège and Peters (2019, 2020) also included collocations with MI scores lower than 3. Third, apart from collocations, Puimège and Peters (2019, 2020) also included other kinds of multiword combinations as the target items.

### *Comprehension*

This study did not find comprehension significantly contributed to the learning of collocations in an academic lecture. This contrasts with Pulido's (2007) findings that reading narrative passages led to vocabulary learning. There are several reasons for this conflicting finding. First, Pulido's study primarily aimed to examine the effects of comprehension on vocabulary learning gains, and therefore, recruited participants with different language proficiency (beginner, intermediate, and advanced). The primary aim of the present study, however, was to compare the vocabulary learning across different input modes. The effect of comprehension on learning gains was a secondary aim. To ensure a fair comparison across different input modes, we recruited participants having the same language proficiency (intermediate); therefore, their comprehension of the academic lecture should be similar. In fact, one-way between groups ANOVA analyses revealed no significant difference in the participants' scores on the comprehension test, $F$ (4,133) = 1.10, $p$= 0.36, $\eta^2$ =0.03. The greater degree of homogeneousity of the participants in the present study compared to those in Pulido's study may explain why we did not find a significant effect of comprehension on learning gains while Pulido did. Another reason for the different results might be this study investigated the learning of collocations in academic lectures whereas Pulido focused on the learning of single words in narrative passages. Comprehension might play different roles in the learning of single words and collocations and when different kinds of input are used. Further research is needed.

### *Prior knowledge of general vocabulary*

This study revealed that prior knowledge of general vocabulary did not significantly contribute to the learning of collocations in an academic lecture. This finding is consistent with Dang et al.'s (2021) findings with the viewing condition. Yet expanding on Dang et al.'s study, this study shows that the same pattern was found in the case of reading, listening, reading while listening, and viewing with captions. The findings of the present study; however, contrast with the findings of other studies which found that prior vocabulary knowledge positively affected the learning of single words through reading (e.g., Horst et al., 1998), reading while listening to graded readers (Webb & Chang, 2015a, 2015b) and viewing television programs (e.g., Peters, 2019; Peters & Webb 2018), and the learning of collocations through viewing televisions (Puimège & Peters, 2019, 2020). The conflicting results may be because the present study and Dang et al. (2021) examined academic input while the other studies investigated non-academic input. Prior knowledge of general vocabulary may play a more significant role in the learning of collocations in non-academic input than in academic input because knowledge of general vocabulary might not always support the learning of specialized vocabulary (Hyland & Tse, 2007). Another reason for the insignificant contribution of prior vocabulary knowledge found in the present study may be because the participants in the present study were fairly homogeneous in terms of knowledge of general vocabulary, which made the effect of prior vocabulary knowledge less transparent.

### *Non-verbal elaboration and verbal elaboration*

Despite the insignificant effect of frequency of occurrence, strength of association, comprehension, and prior knowledge of general vocabulary on the learning of collocations in an academic lecture, nonverbal elaboration, verbal elaboration, and type of vocabulary seem to make significant contributions to the learning. This study showed that collocations with

nonverbal elaboration were more likely to be learned than those without nonverbal elaboration in the case of all experimental groups combined and the listening group. To some extent, this finding was consistent with Yang and Sun's (2013) findings that nonverbal elaboration significantly contributed to the learning of single words. This highlights the importance of nonverbal elaboration in incidental vocabulary learning.

Verbal elaboration only significantly affected learning under the viewing with captions condition. Collocations with explicit elaboration were more likely to be learned than those with implicit elaboration. The deeper learners process the input, the better they notice and learn the linguistic patterns in the input (Craik & Lockhart, 1972; Schmidt, 1990). Explicit elaboration makes the connection between the forms and the meaning of the target vocabulary more salient than implicit elaboration (Vidal, 2011); therefore, collocations with explicit elaboration might draw more attention from the participants and were more likely to be learned than those with implicit elaboration. To some degree, the result of this study is in line with Vidal's (2003, 2011) finding that verbal elaboration had a positive impact on incidental learning of single words, but unlike Vidal (2003, 2011), this study did not find significant effect of verbal elaboration in the case of listening. This might be because the present study measured vocabulary knowledge at the form recognition level while Vidal (2003, 2011) measured vocabulary knowledge at the meaning recall level. Verbal elaboration might have a greater contribution to the learning of the meaning of a new lexical item than its form (Kim, 2006).

***Type of vocabulary***

The type of vocabulary only made a significant unique contribution to learning under the listening condition. Academic collocations are more likely to be learned than technical collocations. To some degree, this finding contrasts previous studies (Vidal, 2003, 2011) which

found that technical vocabulary was slightly easier to be learned than academic vocabulary. The conflicting finding could be explained in two ways. First, this study examined collocations while the previous studies focused on single words. The influence of the type of vocabulary may be different between single words and collocations. Second, unlike Vidal (2003, 2011), this study adopted the current approach toward defining specialized vocabulary (Dang, 2020; Dang et al., 2017; Gardner & Davies, 2014; Nation, 2016), and therefore, better distinguished specialized and non-specialized collocations as well as different levels of specialty of collocations (technical, academic, administrative) in the input. Further research is warranted.

The present study has several limitations. First, it only measured knowledge of collocations at the form recognition level and examined one academic lecture, and did not explicitly measure the participants' listening proficiency. Second, the findings related to the effect of the seven factors are exploratory because the data in this study were limited to the collocations appearing in the lecture. Third, because the lecture was from an open course, it is impossible for us to check why the lecturer decided to adopt certain ways of elaboration in his lecture. Fourth, although it is unlikely that the participants had been exposed to the specialized collocations before the experiment, we did not do follow-up checking with the participants to confirm this fact. Fifth, this study did not explicitly measure the participants' listening proficiency. Future studies should measure incidental learning of collocations at various levels of vocabulary knowledge (e.g., meaning recall, meaning recognition) and the effects of other factors on the learning as well as replicating the study with learners of different levels of English proficiency. It could be useful to explore the reasons why lecturers used certain kinds of elaboration in their talks, and examine the learning of collocations through exposure to multiple academic lectures.

This study provides useful implications for EAP teachers, learners, and content lecturers. It suggests that academic lectures are valuable L2 input for EAP learners and those studying at English medium universities. If academic lectures are used as meaning-focused input for these learners, viewing and viewing with captions are better input modes than listening only, reading only, and reading while listening. Unlike listening only and reading while listening, viewing and viewing with captions lead to collocational learning. Meanwhile, unlike reading only, viewing and viewing with captions offer learners – especially those coming from the educational context with more focus on written input than aural input or those from EFL contexts with limited amounts of aural L2 input – opportunities to be exposed to the aural forms of collocations. The significant effect of verbal elaboration, non-verbal elaboration, and type of vocabulary on the incidental learning of collocations in academic lectures suggests that content lecturers should make use of explicit verbal elaboration and non-verbal elaboration to better facilitate L2 learners' learning of specialized vocabulary in their lectures.

## Conclusion

This study showed that collocations could be learned through reading, viewing, and viewing an academic lecture with captions, and these conditions are likely to make equal contributions to the learning gains. Frequency, the strength of association, comprehension, and prior knowledge of general vocabulary did not significantly affect the learning of collocations in academic lectures, but nonverbal elaboration, type of vocabulary, and type of verbal elaboration did. Overall, this study provides further evidence to support the use of academic lectures for vocabulary learning. It provides us with further insights into the learning of collocations across different input modes and the influence of various factors on learning. It also expanded the multimedia learning theory by providing further evidence supporting the presentation-mode approach, the split-attention

principle, and the redundancy principle. This study is valuable because it is among the very few studies that have investigated the potential of academic lectures for learning collocations, and it is the first to compare the learning of collocations in academic lectures through different input modes.

## Notes

[1] This study does not consider TED Talks as academic lectures.

[2] It is unlikely that the participants had been exposed to the specialized collocations because the present study was conducted at the beginning of the participants' EAP courses, and the participants only studied their discipline-specific courses after having completed the EAP course.

## References


Ayres, P., & Sweller, J. (2014). The split-attention principle in multimedia learning. In R. E. Mayer (Eds.), *Cambridge handbook of multimodal theory* (pp. 206–26). Cambridge: Cambridge University Press.

Biber, D. 1988. *Variation across Speech and Writing*. Cambridge: Cambridge University Press.

Biber, D. 2006. *University Language: A Corpus-Based Study of Spoken and Written Registers*. Amsterdam: John Benjamins.

Biber, D., Conrad, S., & Cortes, V. (2004). If you look at…: Lexical bundles in University teaching and textbooks. *Applied Linguistics*, *25*(3), 371–405. https://doi.org/10.1093/applin/25.3.371

Boers, F., Demecheleer, M., & Eyckmans, J. (2004). Etymological elaboration as a strategy for learning figurative idioms. In P. Bogaards & B. Laufer (Eds.), *Vocabulary in a second language: Selection, acquisition and testing* (pp. 53–78). Amsterdam: John Benjamins.

Boer, F., Lindstromberg, S., & Eyckmans, J (2014). Some explanations for the slow acquisition of L2 collocations. *Vigo International Journal of Applied Linguistics, 11*, 41–62.

Boers, F., & Webb, S. (2018). Teaching and learning collocation in adult second and foreign language learning. *Language Teaching*, *51*(1), 77–89. https://doi.org/10.1017/S0261444817000301

Chaudron, C. (1982). Vocabulary elaboration in teachers' speech to L2 learners. *Studies in Second Language Acquisition, 4*, 170–80.

Coxhead, A. (2018). *Vocabulary and English for Specific Purposes Research: Quantitative and qualitative perspectives*. London: Routledge.

Coxhead, A. (2020). Academic vocabulary. In S. Webb (Ed.), *The Routledge handbook of vocabulary studies* (pp. 97–110). London: Routledge.

Coxhead, A., Dang, T. N. Y., & Mukai, S. (2017). Single and multi-word unit vocabulary in university tutorials and laboratories: Evidence from corpora and textbooks. *Journal of English for Academic Purposes*, *30*, 66–78. https://doi.org/10.1016/j.jeap.2017.11.001

Craik, F., & Lockhart, R.(1972). Levels of processing: A framework for memory research. *Journal of Verbal Learning and Verbal Behavior*, *11*(6), 671–84.

Dang, T. N. Y. (2020). Corpus-based word lists in second language vocabulary research, learning, and teaching. In S. Webb (Eds.), *The Routledge handbook of vocabulary studies* (pp. 288–304). London: Routledge.

Dang, T. N. Y., Coxhead, A., & Webb, S. (2017). The academic spoken word list. *Language Learning*, *67*(4), 959–997. https://doi.org/10.1111/lang.12253

Dang, T. N. Y., Lu, C., & Webb., S (2021). Incidental learning of single words and collocations through viewing an academic lecture. *Studies in Second Language Acquisition, First review*, 1–29. https://doi.org/10.1017/S0272263121000474.

Ellis, R. (1999). *SLA research and language teaching*. Oxford: Oxford University Press.

Feng, M. (2021). *Language Learning through Captioned Videos*. London: Routledge.

Field, J. (2011). Into the mind of the academic listener. *Jounal of English for Academic Purposes*, *10*, 102–112. https://doi.org/10.1016/j.jeap.2011.04.002

Flowerdew, J. (1994). *Academic listening: Research perspectives*. Cambridge: Cambridge University Press.

Gablasova, D., Brezina, V., & McEnery, T. (2017). Collocations in corpus-based language learning research: Identifying, comparing, and interpreting the evidence. *Language Learning*, *67*(S1), 155–179. https://doi.org/10.1111/lang.12225

Gardner, D., & Davies, M. (2014). A new academic vocabulary list. *Applied Linguistics* 35(3), 305–27. https://doi.org/10.1093/applin/amt015

Goh, C. C. M. (2000). A cognitive perspective on language learners' listening comprehension problems. *System*, *28*(1), 55–75. https://doi.org/10.1016/S0346-251X(99)00060-3

Godfroid, A., Ahn, J., Choi, I., Ballard, L., Cui, Y., Johnston, S., … Yoon, H.-J. (2018). Incidental vocabulary learning in a natural reading context: An eye-tracking study. *Bilingualism: Language and Cognition*, *21*(3), 563–584. https://doi.org/10.1017/S1366728917000219

Horst, M., Cobb, T., & Meara, P. (1998). Beyond a clockwork orange: Acquiring second language vocabulary through reading. *Reading in a Foreign Language*, *11*(2), 207–223.

Hyland, K., & Tse, P. (2007). Is there an 'academic vocabulary'? *TESOL Quarterly, 41*(2), 235–53. https://doi.org/10.1002/j.1545-7249.2007.tb00058.x

Jin, Z., & Webb, S. (2020). Incidental vocabulary learning through listening to teacher talk. *Modern Language Journal*, *104*(3), 550–566. https://doi.org/10.1111/modl.12661

Kalyuga, S., & Sweller, S. (2014). The redundancy principle in multimedia learning. In R. E. Mayer (Eds.) *The Cambridge handbook of multimedia learning* (pp. 247–62). Cambridge: Cambridge University Press.

Kim, Y. (2006). Effects of input elaboration on vocabulary acquisition through reading by Korean learners of English as a foreign language. *TESOL Quarterly*, *40*(2), 341–373. https://doi.org/10.2307/40264526

Krashen, S. (1985). *The Input Hypothesis: Issues and Implications.* London: Longman.

Lynch, T. (2011). Academic listening in the 21st century: Reviewing a decade of research'. *Journal of English for Academic Purposes*, *10*(2), 79–88.

Mayer, R. E. 2014. 'Introduction to multimedia learning'. In R. E. Mayer (Eds.) *The Cambridge handbook of multimedia learning* (pp.1–24 ). Cambridge: Cambridge University Press.

Montero Perez, M. (2020). Multimodal input in SLA research. *Studies in Second Language Acquisition, 62*, 653–63.

Nation, I. S. P. (2013). *Learning vocabulary in another language* (2nd ed.). Cambridge: Cambridge University Press.

Nation, I. S. P. (2016). *Making and using word lists for language learning and testing*. Amsterdam: John Benjamins.

Nguyen, T. M. H., & Webb, S. (2017). Examining second language receptive knowledge of collocation and factors that affect learning. *Language Teaching Research*, *31*(3), 298–320. https://doi.org/10.1177/1362168816639619

Paivio, A. (1986). Mental representations: A dual coding approach. Oxford: Oxford University Press.

Pavia, N., Webb, S., & Faez, F. (2019). Incidental vocabulary learning through listening to songs. *Studies in Second Language Acquisition*, *41*(4), 745-768. https://doi.org/10.1017/S0272263119000020

Pellicer-Sánchez, A. (2017). Learning L2 collocations incidentally from reading. *Language Teaching Research*, *21*(3), 381 –402. https://doi.org/10.1177/1362168815618428

Pellicer-Sánchez, A., Tragant, E., Conklin, K., Rodgers, M., Serrano, R., & Llanes, A. (2020). Young learners' processing of multimodal input and its impact on reading comprehension: An eye-tracking study. *Studies in Second Language Acquisition*, *42*, 577–98. https://doi.org/10.1017/S0272263120000091

Peters, E. (2019). The effect of imagery and on-screen text on foreign language vocabulary learning from audiovisual input. *TESOL Quarterly*. https://doi.org/10.1002/tesq.531

Peters, E. (2020). Factors affecting the learning of single-word items. In S. Webb (Eds.), *Routledge handbook of vocabulary studies* (pp. 125–42). London: Routledge.

Peters, E, & Webb, S. (2018). Incidental vocabulary acquisition through watching a single episode of L2 television. *Studies in Second Language Acquisition*, *40*, 551–577. https://doi.org/10.1017/S0272263117000407

Puimège, E., & Peters, E. (2019). Learning L2 vocabulary from audiovisual input: An exploratory study into incidental learning of single words and formulaic sequences.

*Language Learning Journal*, *47*, 424–438. https://doi.org/10.1080/09571736.2019.1638630

Puimège, E., & Peters, E. (2020). Learning formulaic sequences through viewing L2 television and factors that affect learning. *Studies in Second Language Acquisition*. https://doi.org/10.1017/S027226311900055X

Pulido, D. (2007). The relationship between text comprehension and second language incidental vocabulary acquisition: A matter of topic familiarity? *Language Learning*, 57(1), 155–99. https://doi.org/10.1111/j.1467-9922.2007.00415.x

Rodgers, M. P. H., & Webb, S. (2020). Incidental vocabulary learning from viewing television. *ITL-International Journal of Applied Linguistics*, 171(2), 191–220. https://doi.org/10.1075/itl.18034.rod

Schmidt, R. (1990). The role of consciousness in second language learning. *Applied Linguistics 11*(1), 17–46. https://doi.org/10.1093/applin/11.2.129

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

https://doi.org/10.1111/j.1540-4781.2011.01146.x

Simpson-Vlach, R., & Ellis, N. (2010). An Academic Formulas List: New methods in phraseology research. *Applied Linguistics*, *31*(4), 487–512. https://doi.org/10.1093/applin/amp058

Smidt, E., & Hegelheimer, V. (2004). Effects of online academic lectures on ESL listening comprehension, incidental vocabulary acquisition, and strategy use. *Computer Assisted Language Learning*, *17*(5), 517–556. https://doi.org/10.1080/0958822042000319692

Sweller, J., van Merrienboer, J. J. G., & Paas, F. G. W. C. (1998). Cognitive architecture and instructional design. *Educational Psychology Review*, *10*, 251–296.

Szudarski, P. (2012). Effects of meaning-and form-focused instruction on the acquisition of verb-noun collocations in L2 English. *Journal of Second Language Teaching & Research*, *1*(2), 3–37.

Szudarski, P., & Carter, R. (2016). The role of input flood and input enhancement in EFL learners' acquisition of collocations. *International Journal of Applied Linguistics*, *2*(26), 245–265. https://doi.org/10.1111/ijal.12092

Teng, F. (2019). The effects of video caption types and advance organizers on incidental L2 collocation learning. *Computers & Education*. doi: https://doi.org/10.1016/j.compedu.2019.103655.

Uchihara, T., Webb, S. & Yanagisawa, A. (2019). The effects of repetition on incidental vocabulary learning: A meta-analysis of correlational studies. *Language Learning 69*(3), 559–99. https://doi.org/10.1111/lang.12343

van Zeeland, H., & Schmitt, N. (2013). Incidental vocabulary acquisition through L2 listening: A dimensions approach. *System*, *41*, 609–624. https://doi.org/10.1016/j.system.2013.07.012

VanPatten, B. (2007). Input processing in adult second language acquisition. In B. VanPatten & J. Williams (Eds.), *Theories in second language acquisition* (pp. 115–135). Mahwah: Erlbaum.

Vidal, K. (2003). Academic listening: A source of vocabulary acquisition? *Applied Linguistics*, *24*(1), 56–89. https://doi.org/10.1093/applin/24.1.56

Vidal, K. (2011). A comparison of the effects of reading and listening on incidental vocabulary acquisition. *Language Learning*, *61*(1), 219–258. https://doi.org/10.1111/j.1467-9922.2010.00593.x

Webb, S. (2015). Extensive viewing: Language learning through watching television. In D. Nunan & J. C. Richards (Eds.), *Language learning beyond the classroom* (pp. 159–168). New York: Routledge.

Webb, S. (2020). Incidental vocabulary learning. In S. Webb (Eds.), *The Routledge handbook of vocabulary studies* (pp. 225–239). London: Routledge.

Webb, S, & Chang, A. C.-S. (2015a). How does prior word knowledge affect vocabulary learning progress in an extensive reading program? *Studies in Second Language Acquisition*, *37*(4), 651–675. https://doi.org/10.1017/S0272263114000606

Webb, S, & Chang, A. C.-S. (2015b). Second language vocabulary learning through extensive reading with audio support: How do frequency and distribution of occurrence affect learning? *Language Teaching Research*, *19*(6), 667–686. https://doi.org/10.1177/1362168814559800

Webb, S., & Chang, A. C.-S. (2020). How does mode of input affect the incidental learning of collocations? *Studies in Second Language Acquisition*, *First review*, 1-22. https://doi.org/10.1017/S0272263120000297

Webb, S., & Nation, I. S. P. (2017). *How vocabulary is learned*. Oxford: Oxford University Press.

Webb, S, Newton, J., & Chang, A. C.-S. (2013). Incidental learning of collocation. *Language Learning*, *63*, 91–120. https://doi.org/10.1111/j.1467-9922.2012.00729.x

Webb, S., Sasao, Y., & Ballance, O. (2017). The updated Vocabulary Levels Test. *ITL - International Journal of Applied Linguistics*, *168*(1), 33–69. https://doi.org/10.1075/itl.168.1.02web

Yang, H.-C., & Sun, Y.-C. (2013). It is more than knowledge seeking: Examining the effects of OpenCourseWare lectures on vocabulary acquisition in English as a foreign language (EFL) context. *Computer Assisted Language Learning*, *26*(1), 1–20. http://dx.doi.org/10.1080/09588221.2011.624523

## Supporting Information

Additional Supporting Information may be found in the online version of this article at the publisher's website:

**Appendix S1**.Summaries of previous research on incidental learning of collocations

**Appendix S2**. Summaries of previous research on incidental learning of single words in academic lectures

**Appendix S3.** Description of the participants and their learning context

**Appendix S4.** Mean scores (SD) on the Updated Vocabulary Levels Test of the 165 participants

**Appendix S5.** Transcript of the academic lecture

**Appendix S6.** Steps taken when selecting the materials

**Appendix S7.** Target collocations (N= 19)

**Appendix S8.** Steps of selecting the target collocations

**Appendix S9.** Pre-test (English translation)

**Appendix S10.** Comprehension test (English translation)

**Appendix S11.** Steps taken when designing the comprehension test

**Appendix S12.** Steps of the data analysis

**Appendix S13.** Mean scores (SD) on the pretest, immediate posttest, and delayed posttest

**Appendix S14.** Single linear mixed-effects model comparing the scores of the six groups of input modes over the three testing times

**Appendix S15.** Results of the general linear mixed-effects model for all 5 experimental groups combined

**Appendix S16.** Results of the general linear mixed-effects model for the listening group

**Appendix S17.** Results of the general linear mixed-effects model for the viewing with captions group

These appendices are also publicly available at IRIS: https://www.iris-database.org/iris/app/home/detail?id=york:939814

## Acknowledgements

We would like to thank Scott Crossley and the anonymous *Language Learning* reviewers for their constructive feedback, which has greatly improved the quality of this article. We would also like to express our thanks to all learners and expert raters for participating in this study, to Haiyan Gao, Wenhua Liang, Ying Shen for their support with the data collection, and to John Read, Dogus Oksuz, Matt Homer, and Sedar Olmez for their advice on stats.

## The authors

**Thi Ngoc Yen Dang** (first and corresponding author)

School of Education

University of Leeds

LS2 9JT, Leeds, U.K

Email: T.N.Y.Dang@leeds.ac.uk

**Cailing Lu** (second author)

School of Foreign Languages

Shanghai Jiao Tong University

200000, Shanghai, China

Email: lucailing68@126.com

**Stuart Webb** (third author)

Faculty of Education

University of Western Ontario

London, Ontario, Canada

Email: swebb27@uwo.ca

**Thi Ngoc Yen Dang** is a Lecturer at the University of Leeds. Her research interests include vocabulary studies and corpus linguistics. Her articles have been published in *Language*

*Learning, TESOL Quarterly, Language Teaching Research, Studies in Second Language Acquisition, System,* and *Journal of English for Academic Purposes.*

**Cailing Lu** is currently working at Shanghai Jiao Tong University for her postdoctoral research. Her main research interests include vocabulary in English for Specific Purposes and English for Academic Purposes, and corpus linguistics.

**Stuart Webb** is a Professor of Applied Linguistics at the University of Western Ontario. His research interests include vocabulary studies, second language acquisition, and extensive reading, listening, and viewing. His latest books are *How Vocabulary is Learned* (with Paul Nation) and *The Routledge Handbook of Vocabulary Studies*.