

Does writing words in notes contribute to vocabulary learning?

Language Teaching Research

2025, Vol. 29(2) 560–587

© The Author(s) 2024



Article reuse guidelines:

sagepub.com/journals-permissions

DOI: 10.1177/136216882411062184

journals.sagepub.com/home/ltr

Zhouhan Jin  and
Stuart Webb 

University of Western Ontario, Canada

Abstract

There has been little research investigating the effects of notetaking on foreign language (FL) learning, and no studies have examined how it affects vocabulary learning. The present study investigated the vocabulary written in notes of 86 students after they had listened to a teacher in an English as a foreign language (EFL) class. The results showed that 51.2% of participants took notes, and 32.6% wrote information about target words in notes. However, there were only 95 instances of information written about the 28 target words. The results revealed that the odds of vocabulary learning were 15 and 10 times higher in the immediate and delayed posttests for target words that were written in notes. The analysis also indicated that the use of first language (L1) translation in teacher speech increased the chances that target words were written in notes, and that writing words in notes was the most effective predictor of learning.

Keywords

FL vocabulary learning, listening, teacher speech, frequency of occurrence, L1 translation, notetaking, word length

1 Introduction

In natural teaching contexts, the act of notetaking is common (van der Meer, 2012). Students are often encouraged to take notes (Webb & Piasecki, 2018), and even though they are not required to do so, most students record some information delivered by the instructor (Hartley & Davies, 1978). This seems to indicate that both teachers and learners acknowledge the importance of taking notes. Moreover, note taking in lectures is widely accepted because it helps learners to augment their attention and enhance their retention of academic discourse (Machida et al., 2018; Tsai & Wu, 2010).

Corresponding author:

Zhouhan Jin, Graduate Program in Education, University of Western Ontario, 1137 Western Road, London, Ontario, N6G 1G7, Canada

Email: zjin65@uwo.ca

Taking notes is also believed to contribute to foreign language (FL) learning (Oxford, 1990). Writing information in notes may contribute to FL vocabulary learning in two ways. First, the act of writing information in notes may help learners to link what was heard, the spoken forms of words, to the written forms of words. This process of encoding information in notes has been found to contribute to learning (Carrier, 1983; Gharravi, 2018). Second, the information written in notes can subsequently be reviewed to better consolidate learning. Research has also shown that the process of reviewing notes leads to learning (Carrier, 1983; Siegel, 2020).

There are few studies investigating the effects of notetaking on FL learning and no studies that have looked at its effects on vocabulary learning. This is surprising because unfamiliar vocabulary is frequently written in notes in language learning classes (Webb & Piasecki, 2018). There is much that could be learned about the effects of notetaking on vocabulary learning. For example, the extent to which unfamiliar words are written in notes is unknown as is the information that is written about unfamiliar words in notes. Students might write the forms of words with and without their meanings. They might also write words together with their collocates or write words in the phrases in which they were encountered. Understanding the types of lexical information included in notes and the relationship that notetaking has with vocabulary learning may reveal the benefits of notetaking for vocabulary learning, as well as some direction for which types of lexical information students should be encouraged to include in their notes.

There is relatively little research that has examined whether the act of writing words contributes to learning. Several studies have looked at whether writing words in intentional vocabulary learning activities contributes to learning. Thomas and Dieter (1987), Webb and Piasecki (2018), and Candry et al. (2017) found that word writing contributed to FL vocabulary learning, whereas Barcroft (2006) found a negative effect of word writing on vocabulary learning. However, these studies investigated writing words in carefully controlled conditions. In real classroom settings, students may not be instructed to take notes (Hartley & Davies, 1978), and if they do take notes, the students rather than instructors decide on the words, collocations, phrases, sentences, and meanings that they write. Therefore, while the studies exploring the effects of word writing in intentional vocabulary learning activities are useful, they are unlikely to reflect gains that may occur through writing lexical information in notes.

The present study was designed to shed light on the extent to which English as a foreign language (EFL) learners write about unfamiliar words in notes, and the effects of notetaking on vocabulary learning. In addition, three factors (first language translation, frequency of occurrence, word length), which are relevant to the spoken input that EFL learners encounter, were examined to determine whether these factors predict notetaking behavior and contribute to vocabulary learning.

1 Background

Notetaking involves writing down the main ideas and specific information that is presented. It is often regarded as an effective way to facilitate language learning and retention of the language features that are presented to students (Oxford, 1990). Notetaking has two main functions. The first is the encoding function, which can facilitate learning

and retention through the process of listening to oral input and then transferring it from aural to written form (Carrier, 1983; Di Vesta & Gray, 1972). The second is the external storage function, which contributes to learning and retention through later review of the notes (Carrier, 1983; Siegel, 2020). Research on distributed practice, or spacing effects, indicates that the review of information at different timings has a positive effect on learning (e.g. Bahrck, 1979; Ebbinghaus, 1985/1913) and FL vocabulary learning (e.g. Bloom & Shuell, 1981; Nakata & Suzuki, 2019). The positive effect of distributed practice on FL vocabulary learning provides evidence of the value of the external storage function of notetaking for vocabulary learning; if students write unfamiliar words in notes and subsequently study the notes at different timings, this is likely to increase their vocabulary knowledge.

There are no studies that have investigated whether the encoding function of notetaking contributes to vocabulary learning. However, several studies have looked at the effect of word writing on vocabulary learning. These studies provide some indication of whether copying the written form of a word from a visual cue leads to gains in knowledge of form-meaning connection of target items.

Thomas and Dieter (1987) examined the effects of word writing in three experiments in which native English-speaking participants encountered French-English word pairs three times. During the experiments, participants wrote each French word twice in a word writing group while the other group simply viewed the word pairs. Form recall, meaning recall, and free recall tests were used to measure learning after the treatment. The form recall test showed that writing words was most effective when responses were spelled correctly or when they were partially correct (one incorrect letter). The meaning recall test revealed little difference between the two groups, and the free recall test indicated that writing words contributed to greater learning when scoring required correct spelling, and correct spelling except for one letter. Thomas and Dieter concluded that writing words may help to develop knowledge of word forms because it focuses learners on the spellings of words. In contrast, Barcroft (2006) found that word writing inhibited vocabulary learning. In a within-participants design, he compared two paired associate learning conditions. One condition involved viewing word-picture pairs and then writing target words that were shown in each picture and the second condition involved simply viewing the word picture pairs. The results indicated that word writing negatively affected learning gains, which supports Barcroft's (2002) type of processing-resource allocation (TOPRA) model. TOPRA suggests that participants did not have the processing resources to learn both the written forms and their meanings in the word writing condition. In a partial replication of Barcroft (2006), Webb and Piasecki (2018) found that if learners were given sufficient time to write words and direct their attention to form-meaning connection, word writing is likely to contribute to vocabulary learning. This finding provided support for Kroll and Stewart's (1994) Revised Hierarchical Model (RHM) in that associating FL words and their meanings is likely to strengthen their semantic connections. Candry et al. (2017) also found a positive effect for word writing. They compared word-writing and meaning-inferencing conditions. In the word-writing condition, participants were given 30 seconds and required to read sentences containing target words and write the word repeatedly to learn its spelling. In the meaning-inferencing condition, participants were presented with the sentences but rather than

writing the target items, they tried to infer the meanings of the words. The results showed that word writing generated higher scores compared to the meaning-inferencing condition in both form recall and meaning recall. Together, these studies provide some evidence that writing a word may contribute to vocabulary learning. However, there are several important differences between writing words in these studies and writing words in notes during a lesson. In the former, participants are instructed to write words while in the latter, the vocabulary that is included in notes is decided upon by students. It is well documented that intentional vocabulary learning activities tend to yield relatively large gains (Nation, 2013; Webb et al., 2020), and so we should expect word writing in the earlier studies to contribute to learning. In addition, cues in the earlier studies were visual while notetaking in the classroom typically involves writing what was heard in oral input. Research has also indicated that the mode of input can also have an effect on FL vocabulary learning (Brown et al., 2008; Webb & Chang, 2020) with activities involving listening and speaking contributing to smaller gains than those involving reading and writing (Yanagisawa & Webb, 2021).

In the EFL classroom setting, listening to teachers might often be the main source of oral input (Macaro & Tian, 2015; Meunier, 2012). Students may have no access to learn written form through listening and have little opportunity to write down the same words repeatedly because of time pressure (Siegel, 2016). Moreover, notetaking will typically be spontaneous behavior. Most students do take notes for their own use, but they do not necessarily know how to take effective notes (Hartley & Davies, 1978; Piolat et al., 2005). Key information and target vocabulary will not always be presented directly to learners, so they must predict the value of information and attend to key words selectively (James et al., 1988; Kırkgöz, 2010). Research indicates that students' self-selection of vocabulary to learn may lead to greater learning than target vocabulary that has been selected by external agents such as teachers (Laufer & Hulstijn, 2001). No studies to date have looked at the effect of writing words in notes on vocabulary learning in a real classroom setting. This study aims to fill this gap.

2 Factors that contribute to vocabulary learning

Many factors may contribute to learning words encountered in spoken input. Research indicates that increased frequency of occurrence of unfamiliar words has a positive effect on vocabulary acquisition (e.g. Peters & Webb, 2018; Uchihara et al., 2019; Vidal, 2011), and different modes of input may require different numbers of encounters for vocabulary learning to occur (van Zeeland & Schmitt, 2013; Vidal, 2011). Vidal (2011) found that more repeated encounters were needed for vocabulary acquisition through listening than reading. Two and three encounters were sufficient to learn words during reading, but learners needed to encounter words five and six times during listening (Vidal, 2011). van Zeeland and Schmitt (2013) found that it took more than 15 encounters to fully develop and retain word knowledge and suggested that listeners may need to encounter words 50 to 100 times to develop the ability to recall a word's meaning. It stands to reason that the more times unfamiliar words are heard in the classroom, the more likely that learners will attend to these items and write them in notes. However, the relationship between frequency of occurrence, notetaking and vocabulary learning remains to be explored.

Another factor that may contribute to whether words are written in notes is teacher explanation of word meanings. VanPatten's (1996) Input Processing theory emphasized that learners tend to first process word meanings before forms when encoding the input, so teacher explanation of word meanings might focus learners' attention and increase the likelihood that words are written in notes. Zhao and Macaro (2016) investigated the relationship between first language (L1) translation and second language (L2)-only explanations on word meanings and vocabulary acquisition. They found that teacher explanations of word meanings facilitated word learning, and L1 definitions yielded greater gains than L2-only explanations. Lee and Levine (2020) provided additional evidence of the value of elaboration of word meanings in speech. They found that both intermediate and advanced EFL learners who heard L1 translation of target words during listening tasks acquired more vocabulary than learners who did not receive L1 translation. However, participants in both of these studies were not permitted to take notes. In order to make learning more ecologically valid, Jin and Webb (2020) allowed learners to take notes when listening to teacher talk and found a positive effect of L1 translation on vocabulary learning. However, the degree to which L1 translation contributes to notetaking has yet to be examined.

A third factor that may contribute to vocabulary learning through listening is word length. Most research found a negative relationship between word length and word learning (e.g. Barcroft & Rott, 2010; Ellis & Beaton, 1993). However, Puimège and Peters (2019) found that longer words encountered in aural input were more likely to be learnt. Uchihara (2020) also found that the spoken forms of longer words tended to be faster to process than shorter words. They suggest that this is because longer words provide listeners with more cues that aid recognition than shorter words. How the length of words encountered in teacher speech contributes to notetaking and vocabulary learning remains to be examined. A secondary aim of the present study was to examine the relationships between word length, frequency of encounters, L1 translation of target words, and notetaking and vocabulary learning. Examining the extent that different factors affect whether unfamiliar words are written in notes and subsequently learned should provide some indication of the aspects of teacher speech that contribute to notetaking, and in turn, vocabulary learning.

II Research questions and predictions

The following research questions were addressed in the present study:

1. What information do EFL learners write about unknown vocabulary in notes?
2. Does writing unknown target words in notes contribute to vocabulary learning?
3. To what extent do frequency of occurrence, L1 translation, and word length predict whether unknown target words will be written in notes?
4. To what extent do frequency of occurrence, L1 translation, word length, and writing target words in notes predict vocabulary learning?

What information do EFL learners write about unknown vocabulary in notes? Nation (2013) provided an overview of different aspects of word knowledge, including form, meaning, and use. Because the participants encountered FL single words and

collocations in context, and some L1 translations were presented, we predict that FL word form, FL form in FL context, L1 meaning, and L1 meaning with FL form will be included in notes.

Does writing unknown target words in notes contribute to vocabulary learning? Our prediction is that writing unknown target words in notes will contribute to vocabulary learning. This prediction is supported by the Encoding Hypothesis, which emphasizes the positive effect of encoding information on language learning (Carrier, 1983; Di Vesta & Gray, 1972). In addition, several studies (e.g. Candry et al., 2017; Thomas & Dieter, 1987) have indicated that writing words contributes to vocabulary gains.

To what extent do frequency of occurrence, L1 translation, and word length predict whether unknown target words will be written in notes? Learners are more likely to focus their attention on salient or noticeable words. Repeated encounters of words may lead to greater attention paid to those words (Webb & Chang, 2015), and providing L1 translations for unknown words might also make them more salient (Jin & Webb, 2020). Therefore, it is hypothesized that frequency of occurrence and L1 translation predict target words that will be written in notes. Although no studies have looked at the relationship between word length and notetaking, we predict that the shorter a word is, the more likely that it will be written in notes. Teacher speech puts notetakers under pressure that requires them to listen, comprehend, select, and write down words simultaneously (Piolat et al., 2005), so they might be more likely to write down shorter words due to time constraints.

To what extent do frequency of occurrence, L1 translation, word length, and writing target words in notes predict vocabulary learning? Several studies have indicated that L1 translation (e.g. Jin & Webb, 2020; Zhao & Macaro, 2016) and writing words (e.g. Candry et al., 2017; Thomas & Dieter, 1987) contribute to vocabulary learning, so L1 translation and writing target words in notes are predicted to positively affect vocabulary learning. Word length, on the other hand, may also predict vocabulary learning because most research found a negative relationship between word length and word learning (e.g. Barcroft & Rott, 2010; Willis & Ohashi, 2012). It can be expected that the shorter a word is, the more likely it will be learned. However, we hypothesize that frequency of occurrence will not predict word learning because listeners may need to encounter unknown single words 10 times (Webb, 2007) and collocations 15 times (Webb et al., 2013) for sizable learning gains to occur. However, most target words in this study do not reach a sufficient frequency of occurrence for substantial learning to take place.

III Methods

I Participants

A power analysis was conducted to estimate the minimum total sample using G*Power software (Faul et al., 2007) with the following input data: Effect size (f) = medium (0.25), statistical significance level (α) = .05, power ($1 - \beta$) = 0.8, number of groups = 2, number of repeated measures = 3, moderate correlation among repeated measures = 0.5 and nonsphericity correction (ϵ) = 0.75. A total sample of 34 is needed. The original participant pool ($N = 96$) was drawn from students majoring in English at a university in China. Because 10 participants missed one or more of the sessions, the final participant pool of 86 students were included in this study, the size of which was adequate for the

Table 1. Means (*M*s), standard deviations (*SD*s), and 95% confidence intervals (*CI*s) of the word frequency levels on the UVLT.

1,000		2,000		3,000		4,000		5,000	
<i>M</i> (<i>SD</i>)	95% <i>CI</i>	<i>M</i> (<i>SD</i>)	95% <i>CI</i>	<i>M</i> (<i>SD</i>)	95% <i>CI</i>	<i>M</i> (<i>SD</i>)	95% <i>CI</i>	<i>M</i> (<i>SD</i>)	95% <i>CI</i>
29.67 (1.14)	[29.43, 29.92]	28.36 (1.83)	[27.97, 28.75]	25.22 (3.35)	[24.50, 25.94]	18.85 (4.89)	[17.80, 19.90]	14.70 (5.37)	[13.55, 15.85]

Notes. Maximum score = 30. Standard deviations are in parentheses. $N = 86$.

experiment. However, group size cannot be controlled because notetaking in this study was not a compulsory behavior. Forty-four of the 86 participants took notes, and 28 of them wrote target words in notes. Because the goal of this study was to examine the effect of writing unknown words on vocabulary learning, two groups (28 participants who wrote target words in notes and 58 who did not) were included later in the analysis. The study was conducted in accordance with ethical guidelines and with the approval of the institutional ethical review board.

All participants had been learning English for a minimum of 10 years and ranged in age from 18 to 21 years old ($M = 18.8$, $SD = 0.75$). They were enrolled in the same English courses receiving approximately 13.5 hours of foreign language instruction each week. The updated vocabulary levels test (UVLT; Webb et al., 2017) was used to assess the participants' prior vocabulary knowledge. UVLT scores are of value in determining generalizability because the UVLT is a widely used instrument that has gone through a great deal of development and validation. It assesses knowledge of words sampled from each of the first five 1,000-word frequency levels which account for the greatest proportion of spoken and written English (Webb et al., 2017). Although a study that includes aural input might ideally use a vocabulary test measuring knowledge of spoken form, previous studies have shown that the use of a written vocabulary test did not compromise the findings regarding a strong correlation between the spoken and written aspects of vocabulary (e.g. Milton & Hopkins, 2006; Uchihara & Harada, 2018). The participants achieved average raw scores of 29.6 and 28.4 out of 30 at the first and second 1,000-word levels, respectively. Table 1 shows the results of the UVLT. Moreover, an independent t -test was used to look at whether participants who took notes and those who did not take notes had significant differences in UVLT scores. Levene's test revealed that homogeneity of variances was met ($p = .774$), indicating that the participants' vocabulary levels were equivalent between the two groups.

2 Procedure

In the first week, all participants were told that the purpose of this study was to investigate listening comprehension. This was to ensure that the participants' focus was on the content of the teacher's anecdotes rather than vocabulary learning. They were also informed of the research procedure and participants completed a consent form prior to voluntarily taking part in the study. Then, they were given as much time as needed to complete the UVLT followed by a pretest consisting of a meaning-recall test and a

Table 2. Overview of research procedure.

Week 1	Week 2	Week 3
Consent form	Treatment	Delayed posttest
UVLT	Immediate posttest	(Meaning recall and multiple-
Pretest	(Meaning recall and multiple-	choice test)
(Meaning recall and multiple-	choice test)	Debriefing
choice test)		

Note. UVLT = updated vocabulary levels test.

multiple-choice test. One week later, all participants completed the treatment and an immediate posttest. One week after the treatment, all participants took the same tests again with test items in a different order on the delayed posttest. After the delayed posttest, all participants received a debriefing to clarify the real purpose of this study. The overview of research procedure can be seen in Table 2.

3 Treatment

Before the treatment, each of the participant was provided with a blank paper and told by the instructor that taking notes was not compulsory. All of the sheets were collected after the treatment regardless of whether or not participants took notes. The participants did not know that their notes would be examined, and they were not aware of subsequent vocabulary tests.

The treatment lasted 40 minutes. It consisted of 26 minutes of video-taped teacher's anecdotes and approximately 14 minutes of teacher-student interaction (for video script, see Online Supplementary Material Appendix 1). It is important to note that the participants watched short 4–6-minute segments of the video rather than a 26-minute monologue. Teacher-student interaction was distributed between each segment. Each time the video was stopped during the class, the instructor raised questions related to the topic, so that teacher-student interaction occurred before, during, and after watching the video. The teacher-student interaction consisted of initiation-response-feedback sequences which is a common instructional practice found in EFL classrooms in China (Cullen, 1998). It started with open questions, followed by student discussion and responses, and the instructor's feedback. The questions were easy to discuss and were related to the topics in the video. However, the questions, as well as the teacher's gestures were not related to any of the target words and were unlikely to elicit any target vocabulary. Therefore, the treatment was designed to include the target items in the video-taped teacher's anecdotes, but not in the teacher-student interaction.

4 Materials

a Video-taped material. A video-taped English language talk was given by a non-native speaker who had one-year of experience teaching English in China at the university level. It was used to ensure that word-related variables were precisely controlled. The English language talk related to the instructor's life and study experiences in both China

and Canada. The content was easy to comprehend, and the topic was selected to arouse the participants' interest because they had the same major and similar background as the instructor. The instructor spoke in the way that non-native language teachers typically talk to their students in EFL classes at university in China. The most frequent 2,000 word families provided 93.2% lexical coverage of the spoken input. This is within the 90% and 95% coverage estimate suggested by van Zeeland and Schmitt (2013) for which learners are able to understand aural text. In order to make spoken input natural and comprehensible, the teacher in the video used proper speaking pace, L1 translation, visual support and gestures. The rate of speech in the video was 112 tokens per minute. Blau (1990) demonstrated that the speed of speech delivered by native English speakers was approximately 170 tokens per minute. However, 107 words per minute is viewed as 'lecturelike' (Dunkel, 1988). Because the participants were EFL students, who might have found it hard to comprehend English speech when spoken at a nativelike speed, the rate of speech was considered to be appropriate. The video lasted 26 minutes and contained 2,901 running words. However, the presentation of the video was not as a single 26-minute monologue. Instead, it was divided into several short segments of approximately 4–6 minutes in length with teacher–student interaction dispersed in between.

b Target vocabulary. Eighteen single-word items and ten collocations that were included in the video were chosen as the target vocabulary. The target single-word items were words that were less frequent than the most frequent 3,000 words in Nation's (2012) BNC/COCA lists. Half of 18 single words were provided with L1 translation equivalents (Mandarin) immediately after they were used to establish translation links between languages. It is worth noting that the L1 translations were limited to the meanings of target vocabulary and did not entail additional semantic information. The translations that were provided corresponded to L1 glosses in written text; the words were not taught to participants, but rather the meanings were quickly provided to help facilitate comprehension. The L1 translation was only repeated the first two times the target words were encountered because it would be unnatural for teachers to translate an unknown word every time it occurred. For the other half of the target words, no translation was provided for the target word meanings. The target word items were encountered in the teacher's speech 3 to 10 times, providing them the chance to be learned incidentally. Research shows that repeated encounters in context may often contribute to learning the forms and meanings of unfamiliar words (Webb, 2008).

Ten target collocations were made up of seven verb-noun collocations and three adjective-noun collocations. All ten items had a statistical strength of co-occurrence typical of collocations with mutual information scores all above three in COCA. The collocations were comprised of words from the first to third 1,000-word-frequency levels in Nation's (2012) BNC/COCA word lists that the participants were likely to know (e.g. *get approval*, *long term*), and were relatively transparent in meaning except *small talk*. Since words made up of target collocations were familiar to participants, collocations were not provided with L1 translations. In addition, word length in collocation is generally determined by the number of words involved in the sequence (Nation, 2013). The word length of each target collocation was two words. The target item frequency of occurrence ranged from 3 to 6 encounters within the video. A full list of target single words and collocations are in Appendix 1 (in Online Supplementary Material).

Table 3. Using the word *blizzard* as an example for each coding category.

Word category	Examples
FL form	blizzard
FL definition	snowstorm
FL form in context	frequent blizzard in winter
L1 meaning	暴风雪
L1 meaning with FL definition	snow storm 暴风雪
L1 meaning with FL form	blizzard 暴风雪
L1 meaning in context	冬天经常有暴风雪

Notes. FL = foreign language. L1 = first language.

5 Coding of notes

The notes that were collected were carefully examined in order to determine the kinds of information that were written in notes and the frequencies of each type of information. Information about target words included in the notes was classified into seven categories: FL form, FL definition, FL form in FL context, L1 meaning, L1 meaning with FL definition, L1 meaning with FL form, and L1 meaning in context. Table 3 shows examples of each coding category drawn from the word *blizzard*.

6 Assessment

A vocabulary test consisting of meaning recall of target single-word items followed by multiple choice form recognition of target collocations was administered to participants in a pretest, immediate post-test, and delayed post-test to track and compare the learning gains of the participants (for an overview of the procedure, see Table 2). The test items were randomly ordered between test administrations to decrease the possibility that the participants could remember the answers by recognizing the order of words.

a Meaning recall test. In order to measure students' knowledge of form-meaning connections of single-word target items, a meaning recall test was used. This test is a commonly used test format in studies of second language vocabulary acquisition through reading (e.g. Laufer, 2003; Peters et al., 2009), listening (e.g. Brown et al., 2008; van Zeeland & Schmitt, 2013), viewing (e.g. Nguyen & Boers, 2019; Peters & Webb, 2018), and writing (e.g. Laufer, 2003; Pichette et al., 2012). Because the participants were exposed to the aural form of the words during listening to teacher speech, the meaning recall test also presented the items to the participants in spoken form. Piloting indicated that participants were able to demonstrate learning when using this format, and the test scores yielded an acceptable Cronbach's Alpha reliability coefficient of 0.754.

There were 36 items in the test. Half of the items were target words and the other 18 items were high frequency words from the 1,000- and 2,000-word levels that the participants were likely to know (for test format, see Online Supplementary Material Appendix 2). The reason to include high frequency words was to encourage participants to complete the test and take it seriously, and also distract the participants' from

deliberately attending to the target words during the pretest. Analysis of the test data only included the responses for the 18 target items.

During the meaning recall test, participants listened to an oral recording of the items with each word repeated twice. The participants were required to write down anything they could recall about the meaning of each word either in Chinese or English. If they did not know the word meaning, they had the option to check *I don't know*. The answer could be a synonym, an explanation, a paraphrase, or anything else they could use to demonstrate their knowledge. The participants had 30 seconds to answer each question.

In the scoring procedure, the analysis initially included scoring for partially correct answers and completely correct answers. The partially correct response such as responses of 糖 ('sugar'), 酱汁 ('sauce'), and 肉汁 ('meat sauce') for the item *gravy* were awarded 0, 0.5, and 1 point, respectively. However, the analyses revealed that both scoring methods provided similar results; all results that were statistically significant with sensitive scoring were also statistically significant with strict scoring. Therefore, only the results where responses were completely correct will be reported. A score of 1 was awarded for fully correct responses and 0 was awarded for all other responses. The maximum obtainable score was 18.

b Multiple-choice test. In order to measure students' receptive knowledge of collocation, a multiple-choice test was created. The test included 20 items of which 10 were the target collocations and the other 10 were high frequency collocations that the participants were expected to know (for test format, see Online Supplementary Material Appendix 2). The analysis only included responses for the 10 target collocations. All of the test items for the target collocations were taken from Nguyen and Webb's (2017) Receptive Knowledge of Collocation Test. Although the target collocations were relatively transparent in meaning, the test items were still challenging because the distractors were plausible options in Chinese. For example, although *get approval* showed some congruency with L1 collocations, the three distractors *ask*, *make* and *put* are all frequently used with *approval* in Chinese. Thus, to select the correct option, the participants needed to notice the collocations that occurred within the spoken input.

The participants listened to a recording of the test. The node words were orally presented along with the five options (The key, three distractors, and an *I don't know* option). The key and three distractors were randomized across the pretest, immediate post-test, and delayed post-test. All data in this test were scored dichotomously with zero for an incorrect response and one for a correct response. The examples of the target item *small talk* in the pretest, immediate posttest, and delayed posttest are shown below.

The participants hear: 'Number four, talk [1sec.] A, clear talk [2 sec.] B, close talk [2 sec.] C, pretty talk [2 sec] D, small talk [2 sec] (then this sequence is repeated)'

The participant sees on paper in the pretest:

4. A. clear B. close C. pretty D. small I don't know

The participants hear: 'Number one, talk [1sec.] A, close talk [2 sec.] B, pretty talk [2 sec.] C, small talk [2 sec] D, clear talk [2 sec] (then this sequence is repeated)'

The participant sees on paper in the immediate posttest:

1. A. close B. pretty C. small D. clear I don't know

The participants hear: 'Number seven, talk [1sec.] A, small talk [2 sec.] B, close talk [2 sec.] C, clear talk [2 sec] D, pretty talk [2 sec] (then this sequence is repeated)'

The participant sees on paper in the delayed posttest:

1. A. small B. close C. clear D. pretty I don't know

7 Analysis

SPSS (Version 23) was used to analyse the data. If there were any correct FL forms, L1 or FL meanings, or FL forms together with L1 or FL meanings of target words written in notes, it was recorded. The criteria of determining whether a target word was written in notes was in accordance with the instructions and scoring of the meaning recall test because students could recall the meanings of target words in Chinese or English.

For the first research question regarding the information that EFL learners write about unknown vocabulary in notes, descriptive statistics were used to provide an overall picture of the number and percentage of target words that were written in notes. When counting non-target words, both individual words and Chinese characters unrelated to target words were included.

For the second research question examining whether writing unknown target words in notes contributes to vocabulary learning, a repeated measures ANOVA was conducted with test timing as the within-participants variable (pretest vs. posttest vs. delayed), and condition (participants who wrote target words in notes, participants who did not write target words in notes) as the between-participants variable. A follow-up pairwise comparison using the Bonferroni test was carried out to examine the mean difference within each condition at each testing time and the mean difference between the participants who wrote target words in notes and those who did not take notes.

The third research question used logistic regression to look at the extent to which three predictors (frequency of occurrence, L1 translation, word length) affect writing unknown target words in notes. The fourth research question used logistic regression to look at the extent to which four predictors (frequency of occurrence, L1 translation, word length, writing target words in notes) affect vocabulary learning. Whether unknown target words were written in notes was the binary outcome variable in the third research question and score per item was the binary outcome variable in the fourth research question.

IV Results

Research question 1

The descriptive statistics for the information about target words that was included in notes is shown in Table 4. There is no overlap among each word category; each target

Table 4. Target word information in notes.

Word category	Total target words in notes	Number and percentage of participants who wrote about target words ($n = 28$)	Percentage of all participants ($N = 86$)
FL form	40	15 (53.6%)	17.4%
FL definition	2	2 (7.1%)	2.3%
FL form in context	8	8 (28.6%)	9.3%
L1 meaning	12	7 (25%)	8.1%
L1 meaning with FL definition	2	2 (7.1%)	2.3%
L1 meaning with FL form	30	12 (42.9%)	14%
L1 meaning in context	1	1 (3.6%)	1.2%

Note. Twenty-eight of the 86 participants wrote target words in notes, and a participant may record more than one word category in notes. FL = foreign language. L1 = first language.

Table 5. Information of target single word items in notes.

Word category	Number of target single words	Gains in posttest (learning rate)	Gains in delayed posttest (learning rate)
FL form	31	17 (54.8%)	13 (41.9%)
FL definition	2	1 (50%)	0 (0%)
FL form in context	2	2 (100%)	2 (100%)
L1 meaning	11	2 (18.2%)	2 (18.2%)
L1 meaning with FL definition	2	0 (0%)	0 (0%)
L1 meaning with FL form	30	26 (86.7%)	21 (70%)
Total	78	48 (61.5%)	38 (48.7%)

Notes. Gains in posttest refers to a target single word that participants did not know in the pretest but was correct in the immediate posttest. Gains in delayed posttest refers to a target single word that participants did not know in the pretest but was correct in the delayed posttest. FL = foreign language. L1 = first language.

word written in notes was only categorized into one of the word types. Table 5 and Table 6 present information about target single words and collocations respectively with their learning and retention rates.

In answer to the first research question, 44 of the 86 participants (51.2%) took notes, and 28 of the 86 (32.6%) wrote information about target words in their notes. There were 95 instances of information written about target words in notes. Of these 95 instances, more than 80% of the occurrences of information were written about target single word items. The most common types of information written about target single words in notes were FL form and L1 meaning with FL form. Table 5 shows that single words were more likely to be learned and retained if they were written down in L1 meaning with FL form compared with words in FL form. In terms of target collocations, only 17 out of 95 occurrences of information written in notes related to these items, and the most common

Table 6. Information about target collocations in notes.

Word category	Number of target collocations	Gains in posttest (learning rate)	Gains in delayed posttest (learning rate)
FL form	9	3 (33.3%)	3 (33.3%)
FL form in context	6	5 (83.3%)	4 (66.7%)
L1 meaning	1	1 (100%)	1 (100%)
L1 meaning in context	1	0 (0%)	0 (0%)
Total	17	9 (52.9%)	8 (47.1%)

Notes. Gains in posttest refers to a target collocation that participants did not know in the pretest but was correct in the immediate posttest. Gains in delayed posttest refers to a target collocation that participants did not know in the pretest but was correct in the delayed posttest. FL = foreign language. L1 = first language.

information types were FL form and FL form in context. Further information can be found in Tables 4, 5 and 6.

Also, 289 non-target words were included in the notes of the 28 participants who wrote information about target words in their notes and 94 non-target words were included in the notes of the other 16 participants who took notes but did not write target words. If we look at all the participants who took notes, an average of 8.7 non-target words were written, which is 4 times higher than target words that were written in notes.

Research question 2

In response to the second research question, because the unknown target words were made up of both single word items and collocations, the analysis was conducted twice. In preliminary analysis, the assumption of normal distribution was met for all but one of the data subsets (meaning recall pretest [distractors]: skewness value of students who did not take target words in notes was 2.41 ($SE = .31$) and the kurtosis value was 7.44 ($SE = .62$)). Data was minimally winsorized due to the small sample size (Salkind, 2010). One extreme high value was reduced in magnitude to a value that is still at the high end of the distribution but not as extreme, then the assumption of normal distribution was met for all variables. The descriptive statistics of the scores for the test measuring receptive knowledge of single word items and collocations are presented in Table 7 and Table 8.

Repeated Measures ANOVA was conducted to determine whether there was a statistically significant difference in the test scores at different retention intervals (pretest, immediate posttest, and delayed posttest) associated with the condition (writing target words in notes, without writing target words in notes). Test scores were dependent variables and two conditions were independent variables.

For meaning recall test data, the assumption of sphericity was violated, as assessed by Mauchly's test of sphericity, $\chi^2(2) = 13.667, p = .001$. Therefore, the Huynh-Feldt correction was applied ($\epsilon = 0.885$). The ANOVA revealed significant interaction between the effects of test timing and condition on single word learning, $F(1.77, 148.67) = 7.64, p = .001, \eta_p^2 = .083$. The pairwise comparison in the meaning-recall test indicated that

Table 7. Means (*M*s), standard deviations (*SD*s), and 95% confidence intervals (*CI*s) of single word items in meaning recall test.

Condition	Pretest		Immediate posttest		Delayed posttest	
	<i>M</i> (<i>SD</i>)	95% <i>CI</i>	<i>M</i> (<i>SD</i>)	95% <i>CI</i>	<i>M</i> (<i>SD</i>)	95% <i>CI</i>
NT (<i>n</i> = 28)	0.89 (1.40)	[0.35, 1.43]	5.07 (3.20)	[3.83, 6.31]	3.64 (3.15)	[2.42, 4.97]
Without-NT (<i>n</i> = 58)	1.22 (1.76)	[0.76, 1.69]	3.52 (2.58)	[2.84, 4.20]	2.86 (2.80)	[2.13, 3.60]

Notes. NT = writing target words in notes. Maximum score is 18. Standard deviations are in parentheses.

Table 8. Means (*M*s), standard deviations (*SD*s), and 95% confidence intervals (*CI*s) of collocations in multiple choice test.

Condition	Pretest		Immediate posttest		Delayed posttest	
	<i>M</i> (<i>SD</i>)	95% <i>CI</i>	<i>M</i> (<i>SD</i>)	95% <i>CI</i>	<i>M</i> (<i>SD</i>)	95% <i>CI</i>
NT (<i>n</i> = 28)	6.07 (0.94)	[5.71, 6.44]	7.32 (1.76)	[6.64, 8.01]	6.82 (1.66)	[6.18, 7.46]
Without-NT (<i>n</i> = 58)	5.69 (1.27)	[5.35, 6.02]	6.36 (1.46)	[5.98, 6.75]	6.26 (1.35)	[5.90, 6.61]

Notes. NT = writing target words in notes. Maximum score is 10. Standard deviations are in parentheses.

the participants with or without writing target words in notes both scored significantly higher from pretest to posttest ($p < .001$, $d = .1.69$, 95% *CI* [0.83, 2.56]; $p < .001$, $d = .1.04$, 95% *CI* [0.49, 1.59], respectively) and from pretest to delayed posttest ($p < .001$, $d = .1.13$, 95% *CI* [0.33, 1.93]; $p < .001$, $d = .70$, 95% *CI* [0.17, 1.23], respectively). To investigate the interaction between condition and test timing, the data were subjected to simple effects analysis with Bonferroni correction. The analysis indicated that there was no significant difference in the pretest scores between condition ($p = .386$, $d = .20$, 95% *CI* [-1.09, 0.42]). However, the participants who wrote target single-words scored significantly higher in the posttest than those who did not write target single-words ($p = .018$, $d = .56$, 95% *CI* [0.28, 2.83]). No significant difference was found between condition in the delayed posttest ($p = .248$, $d = .27$, 95% *CI* [-0.56, 2.12]).

For the multiple-choice test data, Mauchly's Test of Sphericity indicated that the assumption of sphericity had not been violated, $\chi^2(2) = 1.134$, $p = .567$. However, no significant interaction was found between test timing and condition, $F(2, 168) = 1.46$, $p = .236$, $\eta_p^2 = .017$.

Research question 3

For the third research question, a logistic regression was carried out to determine the relationship between writing target words in notes and one or three variables (frequency of occurrence, L1 translation and word length for single-word items; frequency of occurrence for collocations). The words that the 86 participants already knew before the treatment were excluded. The analysis in total was computed for 1,794 observations,

Table 9. Results of logistic regression including estimates, standard errors (SEs), z-values, p-values, odds ratios (ORs), and 95% confidence intervals (CIs) for writing single word items in notes.

Variable	Estimate	SE	z	p	OR	95% CI
Intercept	-4.46	0.63	-7.03	< .001	0.01	[0.00, 0.04]
FoO	0.04	0.06	0.80	.42	1.05	[0.94, 1.17]
L1	1.25	0.31	4.04	< .001	3.49	[1.90, 6.40]
Length	0.04	0.07	0.55	.58	1.04	[0.90, 1.20]

Notes. FoO = frequency of occurrence. CI = confidence interval for odds ratio (OR).

indicating that there were 1,794 items that the participants did not know before the treatment and could potentially be learned. Among the 1,794 observations, 1,434 were for single words and 360 were for collocations. Table 9 presents the results of the logistic regression for writing target single words in notes.

The analysis revealed that only L1 translation is a significant predictor of writing unknown single-word items in notes. The odds ratio for L1 translation ($OR = 3.49$) indicated that when a target word was translated by the teacher, the odds of writing it down was 3.49 times higher than a word that was not translated into L1. In terms of writing down unknown target collocations, frequency of occurrence did not predict that unknown target collocations would be written in notes.

Research question 4

In answer to the fourth research question, a logistic regression was carried out with the data from the immediate and delayed posttests to determine the relationship between learning gains (dependent variable) and frequency of occurrence, L1 translation, word length, and writing target words in notes for single word items; frequency of occurrence and writing target words in notes for collocations. Table 10 presents the results of the logistic regression of the meaning recall test.

The analysis revealed that L1 translation ($p < .001$, $OR = 1.80$), word length ($p < .001$, $OR = 1.14$), and writing target words in notes ($p < .001$, $OR = 15.42$) contributed significantly to the model for single-word items in the immediate posttest. In the delayed posttest, word length ($p < .001$, $OR = 1.16$) and writing target words in notes ($p < .001$, $OR = 10.38$) significantly contributed to word retention. Frequency of occurrence did not contribute significantly to word learning or retention. Writing the target single-word items in notes was the most powerful predictor. When a target word was written in notes, the odds of a correct response were around 15 times higher in the immediate posttest, and around 10 times higher in the delayed posttest. When a target word was translated by the teacher, the odds of a correct response were 1.80 times higher in the immediate posttest but did not significantly affect vocabulary learning in the delayed posttest. It is interesting to note that word length also has a positive relationship with word learning. For each one-letter increase in the length of a word, the odds of a correct response increased by 14% ($OR = 1.14$) in the immediate posttest and by 16% ($OR = 1.16$) in the delayed posttest. In other words, the longer a word was, the more likely it can be learned.

Table 10. Results of logistic regression including estimates, standard errors (SEs), z-values, p-values, odds ratios (ORs), and 95% confidence intervals (CIs) for meaning recall test.

Variable	Immediate posttest				Delayed posttest							
	Estimate	SE	z	p	OR	95% CI	Estimate	SE	z	p	OR	95% CI
Intercept	-2.98	0.33	-9.01	<.001	0.05	[0.03, 0.10]	-3.45	0.38	-9.08	<.001	0.03	[0.02, 0.07]
FoO	0.03	0.03	1.08	.28	1.03	[0.97, 1.10]	0.05	0.03	1.57	.12	1.06	[0.99, 1.13]
LI	0.59	0.14	4.09	<.001	1.80	[1.36, 2.39]	0.24	0.16	1.51	.13	1.28	[0.93, 1.75]
Length	0.13	0.21	3.49	<.001	1.14	[1.06, 1.23]	0.15	0.04	3.55	<.001	1.16	[1.07, 1.27]
NT	2.74	0.32	8.62	<.001	15.42	[8.28, 28.72]	2.34	0.28	8.38	<.001	10.38	[6.01, 17.95]

Notes. FoO = frequency of occurrence. LI = first language. NT = writing target words in notes. CI = confidence interval for odds ratio (OR).

For learning collocations in the posttest and delayed posttest, neither frequency of occurrence nor writing target words in notes had significant correlations with scores from the immediate and delayed posttests, indicating that both variables did not predict that collocations would be learned.

V Discussion

Although notetaking is considered a pervasive and important activity in classroom contexts, no studies to date have looked at its effect on vocabulary learning in a real classroom setting. This study filled this gap by investigating the information that EFL learners tend to write in notes and whether writing unknown words in notes contributes to vocabulary learning. The present study also examined whether frequency of occurrence, L1 translation, and word length can predict notetaking of target single words, and whether frequency of occurrence can predict notetaking of target collocations. A secondary aim of this study was to include writing unknown words in notes as one more predictor to examine the relationship between those predictors and the vocabulary learning of unknown single words and collocations, respectively.

1 Information about unknown vocabulary in notes

In answer to the first research question, the results showed that more than half of the participants took notes (44 out of 86). This is not surprising because in most teaching and learning contexts, the act of notetaking is common (van der Meer, 2012; Williams & Eggert, 2002). Students are often encouraged to write down novel words (Webb & Piasecki, 2018), and even if they are not explicitly instructed to do so, students are also likely to record lecture content (Williams & Eggert, 2002). However, relatively few target words were written in notes. If every participant wrote every target word, there would be 2,436 occurrences of information written in notes, but the result showed only 95 instances of information about target words were written in notes. This suggests that writing unknown words in notes when listening to teacher speech is uncommon.

There are several reasons that might explain why few unknown words were written in notes. The small number of target words written in notes may be because of their focus on understanding the content, as well as the fact that EFL learners often have difficulty taking notes while listening because it may distract them from attending to subsequent information necessary to understand the spoken input (Piolat et al., 2005). This is in line with Barcroft's (2002) TOPRA model, which suggested that learners' form processing and form learning will decrease if their semantically oriented learning increases. Another possible reason may be due to students' lack of knowledge of their written forms as well as their meanings. It is widely accepted that vocabulary knowledge correlates significantly with learners' writing performance (Lin, 2015; Shi & Qian, 2012). Knowledge of word form (e.g. spelling) is required for writing. Although learners may recognize a word's spoken form through listening, they may have difficulty linking its spoken form to its written form if they lack knowledge of spelling rules, especially in the English orthographic system where a grapheme can represent more than one phoneme. Thus, they may choose to avoid writing words that are hard to spell (Nation, 2013). Students may also

struggle with word meanings. Research shows that words are difficult to infer from context (Webb, 2007). If students encounter unknown words and cannot infer their meanings, they might be less likely to write them down.

It is important to consider why some words were written in notes while others were not. Saliency may play a role in determining learners' attention during listening (Vidal, 2011; Webb, 2020). Only 17 out of 95 occurrences of information in notes were for target collocations. Learners might have been less likely to attend to the unknown target collocations because the words that made up target collocations were familiar to participants. This may have led them to ignore collocations when encountered in the input due to their semantic transparency (Laufer & Waldman, 2011). When looking at target single words written in notes, the chance of writing down words that were translated into L1 in the teacher speech was almost 3.5 times higher than those that were not translated, indicating that the use of L1 could make words more salient and encourage the students to write them down. Vidal (2011) compared learning gains between explicit and implicit elaboration and found that explicit elaboration helped listeners to focus on words and make form-meaning connections. Zhao and Macaro (2016) also found that students learned more words if they were translated into L1 rather than L2.

When looking at the information that was written in notes, the results revealed that the most common type of information written about target items was the FL form on its own. This might be expected, because the FL spoken forms of the items were encountered, while only half of the meanings of items were translated. It indicates that learners are able to use their knowledge of spoken and written forms to enable them to correctly spell words that they have encountered in speech. The second most common type of notetaking relating to target single words was writing the L1 meanings of items together with their FL forms, indicating that learners' attention was also focused on establishing the form-meaning link. Associating meanings to FL forms yielded a higher learning rate in posttest and delayed posttest (see Table 5). This is supported by the RHM which suggests that a strong lexical connection, which requires semantic processing, is established when FL is translated to L1, and hence reinforces semantic connections between FL words and the semantic system (Kroll & Stewart, 1994). When looking at collocations, participants wrote the FL forms of collocations (9 times) and wrote them in context (6 times) but rarely wrote the meanings of collocations (2 times). In addition, associating contexts to FL collocations yielded a higher learning rate (see Table 6), the result of which seems to be in accordance with the RHM. However, the findings should be interpreted with caution due to the small number of collocations written in notes. The difference between the proportion of target single word items and collocations that were written in notes was found to be significant ($p = .008$, $d = .54$) indicating that although students may notice the forms of collocates encountered in speech, they are unlikely to write them in notes to the degree to which they write unfamiliar single word items.

2 Writing unknown target words in notes in vocabulary learning

In response to the second research question, the results showed that the gains made by participants who wrote target single words in notes were significantly greater than those who did not write the items in notes. The participants who wrote unknown target

single-word items gained knowledge of 4.18 (23.2%) and 2.75 (15.3%) items on average in the posttest and delayed posttest, whereas the participants who did not write target words learned 2.30 (12.8%) and 1.64 words (9.1%), respectively. This indicates that learners who write unfamiliar words in notes are more likely to learn those words than those who do not. Because notetaking in this study is a spontaneous behavior, students who wrote down target words could be more motivated than others. Investigating motivation was out of the scope of this study, and as a learning predictor, it is very difficult to describe operationally (Mubeen & Reid, 2014). However, notetaking in this study could be an indicator of motivation because notetaking and motivation are highly correlated with each other (Moos, 2009). According to the 'depth-of-processing' hypothesis (Craik & Lockhart, 1972), retention of information is closely related to how the information is processed. Because writing down target words can increase the learners' engagement and involve generative processing (Armbruster, 2000), this finding may point to the fact that when students encounter unknown single words, taking notes can enhance their learning of those words. This expands on studies of incidental vocabulary learning through reading (e.g. Waring & Takaki, 2003), listening (e.g. van Zeeland & Schmitt, 2013), reading while listening (e.g. Brown et al., 2008), and viewing (Peters & Webb, 2018). Because studies of incidental vocabulary learning are highly controlled, they typically do not allow students to interact with each other and use the language that is encountered in the treatment. The present study looks at learning through listening in a more ecological manner because it allowed students to take notes if they chose to. Therefore, perhaps incidental learning gains may be larger than they have appeared in earlier studies (e.g. Brown et al., 2008; van Zeeland & Schmitt, 2013) because in the majority of studies, learners have not been allowed to write words in notes. Our findings suggest that although the number of unknown words written in notes may be relatively small, the act of notetaking enhances the gains that students can make through listening. It should be noted, however, that in the delayed posttest, the learning gains between participants who wrote and did not write target words in notes were not significant. Bohay et al. (2011) found that participants had more difficulty recalling knowledge of the material one week after the treatment. Therefore, if participants have no chance to further encounter words (e.g. review notes), their memory will decay.

The participants who wrote target collocations increased their knowledge of 1.25 (12.5%) and 0.75 (7.5%) out of 10 target collocations on average in the posttest and delayed posttest, whereas the participants who did not write target words learned on average 0.67 (6.7%) and 0.57 words (5.7%), respectively. Although the participants who wrote collocations in notes slightly outperformed those who did not write target collocations, the interaction between time and notetaking was not significant. The likely reason for this is the lack of collocations written in notes. If processing resources are limited, TOPRA may have a trade-off effect between structural (focus on word form) and semantic (focus on word meaning) elaboration. There were only 17 occurrences of information written about target collocations. Webb et al. (2013) found that learners may need to encounter unknown collocations 15 times to reach a sizable learning gain. However, in this study, collocations ranged from 3 to 6 encounters. Therefore, there may have been an insufficient number of encounters with target collocations to elicit notetaking of the items. Another possible reason why few collocations were written in notes is that

collocations were not presented with L1 translations. As discussed earlier, word saliency affects the probability that words are recorded. Translating collocations may make them more salient and more likely to be written in notes, and hence yield higher learning gains.

3 The relationship between frequency of occurrence, L1 translation, word length and unknown target words written in notes

A positive relationship ($OR = 3.49$) was found between L1 translation and students' writing of target words in this study. This finding indicates that providing the L1 translation of unknown words increases the chances that they will be written in notes. This supports Input Processing (VanPatten, 1996) as learners are more likely to process word meaning before they process it for form, and it is also consistent with the RHM, which suggests that the L1 is likely to play a role in FL processing. This finding helps to explain how the use of L1 translation has a positive effect on word learning. Earlier studies have demonstrated that L1 translation is an efficient way to boost learning in L2 classrooms (e.g. Cook, 2001; Swain & Lapkin, 2000). The results of the present study suggest that it also enhances the potential for vocabulary learning by increasing the likelihood that words will be used productively and written in notes. This in turn provides further support for the use of L1 translation in teacher speech, as it increases the amount of evidence suggesting that L1 translation by teachers positively affects FL vocabulary learning in the classroom (Lee & Levine, 2020; Zhao & Macaro, 2016). However, neither frequency of occurrence nor word length could significantly affect students' writing of target single words. The increased semantic learning may deplete learners' focus on word form as illustrated by TOPRA whereas word length may not show differences in word saliency while processing input as suggested by Carroll (2012). The insignificant relationship between frequency of occurrence and writing target collocations in notes is aligned with the prediction. It might result from the insufficient number of encounters and lack of word saliency as collocations were made up of known words and were encountered 3 to 6 times during the teacher speech.

4 The relationship between frequency of occurrence, L1 translation, word length, writing target words in notes and vocabulary learning

Of the four variables (frequency of occurrence, L1 translation, word length, and writing target words in notes) investigated for learning of single word items, writing target words in notes was the factor with the greatest effect. Many studies highlighted the positive impact that frequency of occurrence has on learning words (e.g. van Zeeland & Schmitt, 2013; Vidal, 2003), and a great deal of research also indicated that L1 translation affect vocabulary learning (e.g. Jin & Webb, 2020; Lee & Levine, 2020; Zhao & Macaro, 2016). However, the results of the present study indicate that writing unknown words in notes has a larger impact on vocabulary learning than the three other variables. This provides support for the Encoding Hypothesis, which emphasized the value of taking notes to better encode and learn information (Carrier, 1983; Di Vesta & Gray, 1972). When students hear an unfamiliar word, writing it down helps them link its spoken form with its written form and brings about a deeper processing of the information.

L1 translation was the second best predictor of the vocabulary gains obtained through listening. Among 78 single words that were written in notes, 62 of them were translated into L1 by the teacher, and the odds of a correct response were 1.80 times higher in the immediate posttest. This suggests that FL words can be learned through taking notes, and the use of L1 translations increases the likelihood that FL words will be learned. However, L1 translation did not lead to a significant effect in the delayed posttest. This does not mean that L1 translation does not have value. There will typically be decay of vocabulary knowledge if words are not encountered further (e.g. Waring & Takaki, 2003). Moreover, it is possible that the effect of L1 translation was mediated by notetaking, because the use of L1 translation by the instructor increased the likelihood that the students would write the translated words in notes.

The results revealed that word length also affected vocabulary learning, and that longer words were more likely to be learned and retained than shorter words. This contrasts our predictions and most studies, which have found negative correlations between word length and word learning (e.g. Barcroft & Rott, 2010; Willis & Ohashi, 2012). However, it is supported by Puimège and Peters (2019) who found that longer words might be more salient in aural input and therefore more likely to be learnt, and Uchihara (2020) found that the spoken forms of longer words tended to be faster to process than shorter words. Because word length has received little attention in vocabulary studies in which the spoken forms of words were encountered, it would be useful to further investigate the contribution of word length to vocabulary learning through listening.

One factor that did not predict vocabulary learning was frequency of occurrence. It may be that frequency of occurrence contributes to increased attention to unknown words and can positively affect vocabulary learning. However, perhaps frequency of occurrence within a small amount of input is not what ultimately determines whether a word is learned. Instead, the results of this study indicate that it may be what happens once attention is drawn to a word that determines whether that word is learned. In this case, it is writing the word in notes that proves more important. It may be different forms of productive use of words such as writing words in notes, discussing the meanings of words with peers, family members, or teachers that has the greatest impact on incidental vocabulary learning (Webb, 2020). It would be useful for future studies to investigate the extent to which different forms of productive use of words that occur after words have been encountered in input affect incidental vocabulary learning.

The analyses also showed that neither frequency of occurrence nor word writing had a significant effect on learning target collocations. This is likely due in part to the small number of encounters with target collocations in the spoken input, as well as the small number of these items that were written in notes. Moreover, because the meanings of most collocations were transparent, and the words that made up the collocations were familiar to learners, they may not recognize these co-occurrences of words as having been encountered in speech. In addition, L1 interference in word combinations may make it difficult to identify FL collocations (Chan & Liou, 2005). For example, *gain knowledge* and *receive knowledge* have the same meaning in Chinese, making it difficult to identify *gain knowledge* as the correct FL collocation without actually noticing its occurrence. Because the present study is the first to look at the contributions of

notetaking to the learning of collocations, it would be useful for future research to explore the extent to which writing FL collocations contributes to learning when learners are exposed to larger quantities of spoken input with a greater number of exposures to items.

VI Pedagogical implications

The results suggest that learners should be encouraged to write unfamiliar words that are encountered in notes. However, it is important to consider whether an explicit focus on notetaking and writing more words in notes would alter the effects found in this study; it is possible that writing a small number of salient items in notes contributes to larger gains than writing a larger number of words at the encouragement of teachers. It is also important for students to consider which words are worth writing as not all words are equally useful (Webb & Nation, 2017). The relatively small number of target words written in notes suggests that students would benefit from learning effective notetaking strategies, which can help listeners to predict information actively and attend to key words selectively (James et al., 1988; Kırkgöz, 2010). For example, teachers could insert brief pauses before the target words, provide sufficient word exposures, and use oral organizational cues in teacher talk to draw learners' attention to useful words and encourage them to record notes efficiently (Titsworth & Kiewra, 2004). The findings of the present study also indicate that translating target words is a useful way to draw students' attention to important vocabulary and increase the chances that the words will be written in notes. One reason why students tend to put effort and faith into notetaking is because notes can be used later while reviewing for subsequent assessment (Kırkgöz, 2010; Siegel, 2020), and some studies demonstrated that students' improved recall is benefited from both taking and reviewing notes (e.g. Fisher & Harris, 1973). However, because of the lack of research on two key variables examined in the research (notetaking and listening to teacher talk), this study did not investigate the utility of later review. Therefore, it would also be useful for future research to compare vocabulary learning through taking notes only and through both taking and reviewing notes to better investigate the effectiveness of notetaking.

In addition, it should be noted that because this study looked at the spontaneous behavior of writing words in notes in a real classroom setting, there was a large unequal sample size which may affect the results. Also, the participants in this study took notes in their own methods. Research shows that students who took notes according to their own method showed lower levels of language achievement than those who received explicit instructions such as using graphic organizers (Rahmani & Sadeghi, 2011) or Cornell Method of notetaking (Hayati & Jalilifar, 2009). Therefore, it would be useful for future research to increase sample size to better generalize findings and to investigate whether instructed notetakers can yield better results on vocabulary learning than uninstructed notetakers through listening to teacher talk.

VII Conclusions

Overall, this study revealed that more than half of the participants took notes in a classroom setting but few unknown words were written in those notes. Results highlighted the

types of lexical information included in notes and that writing words in notes contributed to significant vocabulary learning. This suggests that increasing learners' attention to unknown words encountered in speech and encouraging learners to write unfamiliar vocabulary in notes may be an effective strategy to promote vocabulary learning. This study also revealed a positive relationship between vocabulary learning of individual words and three predictors (writing unknown words in notes, L1 translation, word length), and that writing unknown words in notes was the greatest predictor of learning. Because this was the first study to explore the potential for learning words through note-taking, it would be useful for future studies to explore approaches to notetaking and their relative contributions to vocabulary learning.

Funding

The author(s) received no financial support for the research, authorship, and/or publication of this article.

ORCID iDs

Zhouhan Jin  <https://orcid.org/0000-0001-5285-9384>

Stuart Webb  <https://orcid.org/0000-0002-8297-4997>

Supplemental material

Supplemental material for this article is available online.

References

- Armbruster, B.B. (2000). Taking notes from lectures. In Flippo, R.F., & D.C. Caverly (Eds.), *Handbook of college reading and study strategy research* (pp. 175–199). Lawrence Erlbaum.
- Bahrick, H.P. (1979). Maintenance of knowledge: Questions about memory we forgot to ask. *Journal of Experimental Psychology: General*, *108*, 296–308.
- Barcroft, J. (2002). Semantic and structural elaboration in L2 lexical acquisition. *Language Learning*, *52*, 323–363.
- Barcroft, J. (2006). Can writing a new word detract from learning it? More negative effects of forced output during vocabulary learning. *Second Language Research*, *22*, 487–497.
- Barcroft, J., & Rott, S. (2010). Partial word form learning in the written mode in L2 German and Spanish. *Applied Linguistics*, *31*, 623–650.
- Blau, E.K. (1990). The effect of syntax, speed, and pauses on listening comprehension. *TESOL Quarterly*, *24*, 746–753.
- Bloom, K.C., & Shuell, T.J. (1981). Effects of massed and distributed practice on the learning and retention of second-language vocabulary. *The Journal of Educational Research*, *74*, 245–248.
- Bohay, M., Blakely, D.P., Tamplin, A.K., & Radvansky, G.A. (2011). Note taking, review, memory, and comprehension. *The American Journal of Psychology*, *124*, 63–73.
- Brown, R., Waring, R., & Donkaewbua, S. (2008). Incidental vocabulary acquisition from reading, reading-while-listening, and listening to stories. *Reading in a Foreign Language*, *20*, 136–163.
- Candry, S., Elgort, I., Deconinck, J., & Eyckmans, J. (2017). Word writing vs. meaning inferencing in contextualized L2 vocabulary learning: Assessing the effect of different vocabulary learning strategies. *Canadian Modern Language Review*, *73*, 293–318.

- Carrier, C.A. (1983). Notetaking research: Implications for the classroom. *Journal of Instructional Development*, 6, 19–26.
- Carroll, S.E. (2012). When is input salient? An exploratory study of sentence location and word length effects on input processing. *IRAL – International Review of Applied Linguistics in Language Teaching*, 50, 39–67.
- Chan, T.P., & Liou, H.C. (2005). Effects of web-based concordancing instruction on EFL students' learning of verb–noun collocations. *Computer Assisted Language Learning*, 18, 231–251.
- Cook, V. (2001). Using the first language in the classroom. *Canadian Modern Language Review*, 57, 402–423.
- Craik, F.I.M., & Lockhart, R.S. (1972). Levels of processing: A framework for memory research. *Journal of Verbal Learning and Verbal Behavior*, 11, 671–684.
- Cullen, R. (1998). Teacher talk and the classroom context. *ELT Journal*, 52, 179–187.
- Di Vesta, F.J., & Gray, G.S. (1972). Listening and note taking. *Journal of Educational Psychology*, 63, 8–14.
- Dunkel, P. (1988). The effect of delivery modification on notetaking and comprehension of L2 lectures. *Journal of Intensive English Programs*, 2, 41–52.
- Ebbinghaus, H. (1885/1913). *Memory: A contribution to experimental psychology*. Translated by Ruger, H.A., & Bussenius, C.E. Teachers College Press.
- Ellis, N.C., & Beaton, A. (1993). Psycholinguistic determinants of foreign language vocabulary learning. *Language Learning*, 43, 559–617.
- Faul, F., Erdfelder, E., Lang, A.G., & Buchner, A. (2007). Effects of notetaking instruction on intermediate and advanced L2 English learners: A quasi-experimental study. *Behavior Research Methods*, 39, 175–191.
- Fisher, J.L., & Harris, M.B. (1973). Effect of note taking and review on recall. *Journal of Educational Psychology*, 65, 321–325.
- Gharravi, A.M. (2018). Impact of instructor-provided notes on the learning and exam performance of medical students in an organ system-based medical curriculum. *Advances in Medical Education and Practice*, 9, 665–672.
- Hartley, J., & Davies, I.K. (1978). Note-taking: A critical review. *Innovations in Education and Teaching International*, 15, 207–224.
- Hayati, A.M., & Jalilifar, A. (2009). The impact of cultural knowledge on listening comprehension of EFL learners. *English Language Teaching*, 2, 101–111.
- James, K., Jordan, R.R., & Matthews, A.J. (1988). *Listening comprehension and note-taking course*. Collins.
- Jin, Z., & Webb, S. (2020). Incidental vocabulary learning through listening to teacher talk. *Modern Language Journal*, 104, 550–566.
- Kırkgöz, Y. (2010). Promoting students' note-taking skills through task-based learning. *Procedia – Social and Behavioral Sciences*, 2, 4346–4351.
- Kroll, J.F., & Stewart, E. (1994). Category interference in translation and picture naming: Evidence for asymmetric connections between bilingual memory representations. *Journal of Memory and Language*, 33, 149–174.
- Laufer, B. (2003). Vocabulary acquisition in a second language: Do learners really acquire most vocabulary by reading? Some empirical evidence. *Canadian Modern Language Review*, 59, 567–587.
- Laufer, B., & Hulstijn, J. (2001). Incidental vocabulary acquisition in a second language: The construct of task-induced involvement. *Applied Linguistics*, 22, 1–26.
- Laufer, B., & Waldman, T. (2011). Verb–noun collocations in second language writing: A corpus analysis of learners' English. *Language Learning*, 61, 647–672.

- Lee, J.H., & Levine, G.S. (2020). The effects of instructor language choice on second language vocabulary learning and listening comprehension. *Language Teaching Research, 24*, 250–272.
- Lin, L.H.F. (2015). Role of depth of vocabulary knowledge in EFL Learners' writing proficiency. Unpublished doctoral dissertation, Hong Kong Polytechnic University, Hong Kong.
- Macaro, E., & Tian, L. (2015). Exploring teachers' oral explanations of new English lexical items in a Chinese university: Comparisons with dictionary information. *System, 52*, 78–90.
- Machida, K., Chin, M., & Johnson, K.A. (2018). The provision of partial notes is not associated with improved student attention in lectures or subsequent understanding of the lecture material. *Active Learning in Higher Education, 19*, 101–115.
- Meunier, F. (2012). Formulaic language and language teaching. *Annual Review of Applied Linguistics, 32*, 111–129.
- Milton, J., & Hopkins, N. (2006). Comparing phonological and orthographic vocabulary size: Do vocabulary tests underestimate the knowledge of some learners. *Canadian Modern Language Review, 63*, 127–147.
- Moos, D.C. (2009). Note-taking while learning hypermedia: Cognitive and motivational considerations. *Computers in Human Behavior, 25*, 1120–1128.
- Mubeen, S., & Reid, N. (2014). The measurement of motivation with science students. *European Journal of Educational Research, 3*, 129–144.
- Nakata, T., & Suzuki, Y. (2019). Effects of massing and spacing on the learning of semantically related and unrelated words. *Studies in Second Language Acquisition, 41*, 287–311.
- Nation, I.S.P. (2012). *The BNC/COCA word family lists*. Available at: <http://www.wgtn.ac.nz/lals/about/staff/paul-nation> (accessed November 2021).
- Nation, I.S.P. (2013). *Learning vocabulary in another language*. 2nd edition. Cambridge University Press.
- Nguyen, C.D., & Boers, F. (2019). The effect of content retelling on vocabulary uptake from a TED talk. *TESOL Quarterly, 53*, 5–29.
- Nguyen, T.M.H., & Webb, S. (2017). Examining second language receptive knowledge of collocation and factors that affect learning. *Language Teaching Research, 21*, 298–320.
- Oxford, R.L. (1990). *Language learning strategies: What every teacher should know*. Newbury House.
- Peters, E., & Webb, S. (2018). Incidental vocabulary acquisition through viewing L2 television and factors that affect learning. *Studies in Second Language Acquisition, 40*, 551–577.
- Peters, E., Hulstijn, J.H., Sercu, L., & Lutjeharms, M. (2009). Learning L2 German vocabulary through reading: The effect of three enhancement techniques compared. *Language Learning, 59*, 113–151.
- Pichette, F., De Serres, L., & Lafontaine, M. (2012). Sentence reading and writing for second language vocabulary acquisition. *Applied Linguistics, 33*, 66–82.
- Piolat, A., Olive, T., & Kellogg, R.T. (2005). Cognitive effort during note taking. *Applied Cognitive Psychology, 19*, 291–312.
- 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.
- Rahmani, M., & Sadeghi, K. (2011). Effects of note-taking training on reading comprehension and recall. *The Reading Matrix, 11*, 116–128.
- Salkind, N.J. (2010). Winsorize. In Salkind, N.J. (Ed.), *Encyclopedia of research design: Volume 1* (pp. 1637–1638). Sage.

- Shi, L., & Qian, D. (2012). How does vocabulary knowledge affect Chinese EFL learners' writing quality in web-based settings? Evaluating the relationships among three dimensions of vocabulary knowledge and writing quality. *Chinese Journal of Applied Linguistics*, 35, 117–127.
- Siegel, J. (2016). A pedagogic cycle for EFL note-taking. *ELT Journal*, 70, 275–286.
- Siegel, J. (2020). Effects of notetaking instruction on intermediate and advanced L2 English learners: A quasi-experimental study. *Journal of English for Academic Purposes*, 46, 100868.
- Swain, M., & Lapkin, S. (2000). Task-based second language learning: the uses of the first language. *Language Teaching Research*, 4, 251–274.
- Thomas, M.H., & Dieter, J.N. (1987). The positive effects of writing practice on integration of foreign words in memory. *Journal of Educational Psychology*, 79, 249–253.
- Titsworth, B.S., & Kiewra, K.A. (2004). Spoken organizational lecture cues and student note-taking as facilitators of student learning. *Contemporary Educational Psychology*, 29, 447–461.
- Tsai, T.-F., & Wu, Y. (2010). Effects of note-taking instruction and note-taking languages on college EFL students' listening comprehension. *New Horizons in Education*, 58, 120–132.
- Uchihara, T. (2020). The effects of spoken input on learning the spoken forms of second language words: Studies of frequency of exposure, acoustic variability, and mode of input. Unpublished doctoral dissertation, University of Western Ontario, London, ON, Canada.
- Uchihara, T., & Harada, T. (2018). Roles of vocabulary knowledge for success in English-medium instruction: Self-perceptions and academic outcomes of Japanese undergraduates. *TESOL Quarterly*, 52, 564–587.
- Uchihara, T., Webb, S., & Yanagisawa, A. (2019). The effects of repetition on incidental vocabulary learning: A meta-analysis of correlational studies. *Language Learning*, 69, 559–599.
- van der Meer, J. (2012). Students' note-taking challenges in the twenty-first century: Considerations for teachers and academic staff developers. *Teaching in Higher Education*, 17, 13–23.
- VanPatten, B. (1996). *Input processing and grammar instruction*. Ablex.
- Van Zeeland, H., & Schmitt, N. (2013). Lexical coverage in L1 and L2 listening comprehension: The same or different from reading comprehension? *Applied Linguistics*, 34, 457–479.
- Vidal, K. (2003). Academic listening: A source of vocabulary acquisition? *Applied Linguistics*, 24, 56–89.
- Vidal, K. (2011). A comparison of the effects of reading and listening on incidental vocabulary acquisition. *Language Learning*, 61, 219–258.
- Waring, R., & Takaki, M. (2003). At what rate do learners learn and retain new vocabulary from reading a graded reader? *Reading in a Foreign Language*, 15, 130–163.
- Webb, S. (2007). The effects of repetition on vocabulary knowledge. *Applied Linguistics*, 28, 46–65.
- Webb, S. (2008). The effects of context on incidental vocabulary learning. *Reading in a Foreign Language*, 20, 232–245.
- Webb, S. (2020). Incidental vocabulary learning. In Webb, S. (Ed.), *The Routledge handbook of vocabulary studies* (pp. 225–239). Routledge.
- Webb, S., & Chang, A.C.-S. (2015). Second language vocabulary learning through extensive reading with audio support: How do frequency and distribution of occurrence affect learning? *Language Teaching Research*, 19, 667–686.
- Webb, S., & Chang, A.C.-S. (2020). How does mode of input affect the incidental learning of collocations? *Studies in Second Language Acquisition*. Epub ahead of print 2 September 2020. DOI: 10.1017/S0272263120000297.
- Webb, S., & Nation, I.S.P. (2017). *How vocabulary is learned*. Oxford University Press.
- Webb, S., & Piasecki, A. (2018). Re-examining the effects of word writing on vocabulary learning. *International Journal of Applied Linguistics*, 169, 72–94.

- Webb, S., Newton, J., & Chang, A. (2013). Incidental learning of collocation. *Language Learning*, 63, 91–120.
- Webb, S., Sasao, Y., & Ballance, O. (2017). The updated vocabulary levels test: Developing and validating two new forms of the VLT. *ITL – International Journal of Applied Linguistics*, 168, 33–69.
- Webb, S., Yanagisawa, A., & Uchihara, T. (2020). How effective are intentional vocabulary-learning activities? A meta-analysis. *Modern Language Journal*, 104, 715–738.
- Williams, R.L., & Eggert, A.C. (2002). Notetaking in college classes: Student patterns and instructional strategies. *The Journal of General Education*, 51, 173–199.
- Willis, M., & Ohashi, Y. (2012). A model of L2 vocabulary learning and retention. *The Language Learning Journal*, 40, 125–137.
- Yanagisawa, A., & Webb, S. (2021). Involvement Load Hypothesis Plus: Creating an improved predictive model of incidental vocabulary learning. *Studies in Second Language Acquisition*. Epub ahead of print 21 October 2021. DOI: 10.1017/S0272263121000577.
- Zhao, T., & Macaro, E. (2016). What works better for the learning of concrete and abstract words: teachers' L1 use or L2-only explanations? *International Journal of Applied Linguistics*, 26, 75–98.