The Contribution of Aural Input in Incidental Vocabulary Learning

D. Faisal Abdullah Al-Hamoud
Department of English, College of Education, Majmaah University, Saudi Arabia

The Contribution of Aural Input in Incidental Vocabulary Learning

D. Faisal Abdullah Al-Hamoud

Department of English, College of Education, Majmaah University, Saudi Arabia

Search submission date:19/2/2024
Research acceptance date:26/8/2024

Abstract:

It is becoming increasingly evident that L2 learners acquire new words incidentally through exposure to aural input, which leads to vocabulary development. Despite this, our understanding of incidental vocabulary learning from various aural input modes is still limited. Using an experimental, pretest-treatment-posttest design, this study aimed to explore how different treatments: listening only, reading while listening, and listening with visual enhancement (i.e., colors, pictures, and some highlighted words) contributed to vocabulary learning from a non-fiction text. Three levels of vocabulary gain were evaluated: (a) meaning recognition, (b) form recognition (now known as aural-form matching), and (c) meaning recall. To this end, 90 male school learners of EFL in Saudi Arabia were randomly assigned to three groups (N = 30 in each group). The results demonstrated that vocabulary learning took place through the three input conditions at the three levels of word knowledge. However, listening with visual enhancement led to significantly greater vocabulary learning compared to the listening-only and reading-while-listening modes. The results indicated that there was no significant distinction between listening only and reading while listening conditions. Furthermore, meaning recognition showed the largest gains while meaning recall showed the smallest. This suggests a hierarchy in aural vocabulary learning. This study's findings suggest that listening can be enhanced with visual aids to promote incidental vocabulary learning. Pedagogical implications and directions for future research are addressed.

Keywords: aural input, incidental vocabulary learning, reading while listening, visual input enhancement.

مساهمة المدخلات السمعية في تعلم المفردات العرضية

د. فيصل بن عبدالله الحمود

قسم اللغة الإنجليزية – كلية التربية – جامعة المجمعة

ملخص الدراسة:

أصبح من الواضح بشكل متزايد أن متعلمي اللغة الثانية يكتسبون كلمات جديدة عرضياً من خلال التعرض للمدخلات السمعية؛ مما يؤدي إلى تطوير المفردات. وعلى الرغم من ذلك، فلا يزال فهمنا محدوداً حول تعلم المفردات العرضي عبر أوضاع المدخلات السمعية المختلفة. واستكشفت هذه الدراسة، باستخدام تصميم الاختبار القبلي والبعدي، كيف ساهمت المعالجات المختلفة: الاستماع فقط، والقراءة أثناء الاستماع، والاستماع مع التحسين البصري (أي الألوان والصور وبعض الكلمات المميزة) في تعلم المفردات من نص غير خيالي. وتم تقييم ثلاثة مستويات من اكتساب المفردات: (أ) التعرف على المعنى، (ب) التعرف على الشكل (المعروف الآن باسم مطابقة الشكل السمعي)، و (ج) استدعاء المعنى. وتحقيقا لهذه الغاية، وقُسم تسعون طالباً من متعلمي اللغة الإنجليزية لغة أجنبية في المملكة العربية السعودية عشوائيا على ثلاث مجموعات (العدد = 30 في كل مجموعة). أظهرت النتائج أن تعلم المفردات حدث من خلال أوضاع المدخلات الثلاثة في المستويات الثلاثة لمعرفة الكلمات. ومع ذلك أدى الاستماع مع التحسين البصري للمدخلات إلى تعلم مفردات أكبر بكثير من الاستماع فقط والقراءة أثناء الاستماع. كما أشارت النتائج إلى عدم وجود فروقات إحصائية بين الاستماع فقط والقراءة أثناء الاستماع. علاوة على ذلك، أظهر التعرف على المعنى مكاسب أكثر، بينما أظهر استدعاء المعنى مكاسب أقل؛ مما يشير إلى وجود تسلسل في تعلم المفردات السمعية. كما تشير نتائج هذه الدراسة إلى أنه يمكن تعزيز الاستماع باستخدام الوسائل البصرية لتعزيز تعلم المفردات العرضي. وتناول البحث الآثار والاتجاهات التربوية المستقبلية.

الكلمات المفتاحية: المدخلات السمعية، تعلم المفردات العرضية، القراءة أثناء الاستماع، تحسين المدخلات البصرية

Introduction

It is essential to develop a second language (L2) lexicon for improving general language proficiency (Alderson, 2005). Furthermore, incidental vocabulary learning has been reported to be a crucial issue in fostering vocabulary development in L2 learners (Chen, 2021, Hulstijn, 2001, Nation & Waring, 2020, and Webb & Nation, 2017). Incidental vocabulary learning refers to the learning process that occurs without particular intention to focus on vocabulary (Nation, 2001). Although intentional approaches to learning vocabulary have been suggested to lead to sizable vocabulary gain (Laufer, 2005), a large number of words L2 learners need to perform language tasks come from incidental vocabulary learning adequately (e.g., Akhtar, 2004; Henderson, Devine, Weighall, and Gaskill, 2015; Peters, 2018). While incidental vocabulary learning has received attention from researchers, an abundance of previous research has been devoted to exploring the role of different modes of reading in the incidental gain of L2 vocabulary. Less research has attracted incidental vocabulary learning from listening (Zhang & Graham, 2020). Among the modes of input for an incidental gain of vocabulary from aural input, viewing L2 movies and television (Peters & Webb, 2018; Rodgers & Webb, 2020), listening to songs (e.g., Pavi, Webb, and Faez, 2019), and viewing L2 captioned videos (e.g., Yeldham, 2018) have been the modes that researchers mainly explored. However, studies that compare different modes of aural input with a single cohort of L2 learners are scarce.

The only recent study which has looked at incidental learning from aural input in different conditions in English as a foreign language (EFL) context is Feng and Webb (2020). Feng and Webb compared vocabulary learning from reading, reading while listening, and viewing (without subtitles). Their results revealed that learning has occurred through these three input modes, but no significant difference was found between them. An earlier study by Neuman

and Koskinen (1992) has examined learning from viewing, with and without captions, and reading while listening. Their findings showed that viewing with captions resulted in a better vocabulary gain than reading while listening. However, without further empirical research which compares the extent to which different modes of aural input contribute to vocabulary gain, it is difficult to draw clear conclusions about the value of one mode over another. To this end, the present study aims at contributing to the existing literature by examining the effectiveness of different modes of aural input on L2 incidental vocabulary learning among school learners in an EFL context. Three modes of aural input are investigated in this study: namely, listening only, reading while listening (i.e., reading a plain text), and listening with visual enhancement (i.e., participants listened while viewing the actual book with pictures, colors, and highlighted target words)
Building on the existing research, which has explored the effect of different modes of input on incidental vocabulary learning, this study seeks to advance our knowledge by examining the impact of aural input in three conditions: listening only, reading while listening, and listening with L2 visual enhancement, on incidental vocabulary learning among EFL schoolchildren. Moreover, the aim of this study is to empirically explore the value of each mode to EFL vocabulary learning. Thus, this would help learners, teachers, and textbook writers utilize extensive exposure to aural input condition that best enriches vocabulary development. In recent years, extensive viewing has been advocated by many researchers (e.g. Peters & Webb, 2018; Webb & Nation, 2017). Hence, this needs to be incorporated in textbooks as well as classroom interactions. To the best of my knowledge, no study has compared the differential effect of these three modes with a single cohort of L2 learners.

Literature Review

Incidental learning of vocabulary through listening

Research has shown that aural input can be a valuable source to support L2 incidental vocabulary learning (e.g., Brown, Waring, and Donkaewbua, 2008; van Zeeland & Schmitt, 2013; Vidal, 2003; 2011). Moreover, Webb and Nation (2017) and Nation and Waring (2020) point out that listening can help L2 learners gain knowledge of new spoken forms, learn grammatical functions and collocations of these words, and guess their meanings from the contexts in which they are encountered. However, studies that compare incidental vocabulary learning from listening and reading modes of inputs produce mixed results. For example, Feng and Webb (2020) explore incidental vocabulary gained from reading, listening, and viewing among a group of 76 Chinese learners of EFL. Their results showed that vocabulary gain was observed through the three input modes, and no significant differences were found between them.

On the other hand, other studies have shown that incidental vocabulary learning did occur more from reading than listening. For example, a study by Brown et al. (2008), which investigates vocabulary learning from three input modes, namely, reading, listening, and reading while listening, has shown that reading and reading while listening conditions contributed more to incidental vocabulary learning than listening only. Vidal (2011) explores incidental vocabulary learning through listening and reading academic texts in the same vein. The findings showed that while both input conditions contributed to vocabulary gain, learning from the reading mode was significantly greater than that from listening. However, the small vocabulary gain reported in Brown et al. (2008) and Vidal (2011) might be attributed to the fact that the written modality of measures used may not be sensitive enough to capture

spoken gains. This may have led to this unnoticed learning from the listening-only mode.

Webb and Chang (2022) explored incidental learning of collocations, among 138 Taiwanese college students, from a graded reader over three weeks. The participants were divided into three input mode groups: listening only, reading only, and reading while listening. The participants' knowledge of 17 targets, and two-word collocations, was measured over three phases: a pretest, an immediate posttest, and a 4-week delayed posttest. Webb and Chang used a matching plus meaning recall test to measure the participants' incidental learning and a separate meaning recall test. The results showed the superiority of reading while listening mode over the other two modes. Furthermore, unlike Brown et al. (2008) and Vidal (2011), Webb and Chang (2022) found the listening-only mode to be as effective as the reading-only mode for fostering incidental learning.

Listening only seems to contribute less to incidental vocabulary learning because it seems that greater cognitive demand is placed on listeners due to the required fast processing of the spoken input (Renandya & Farrell, 2011). Another reason might be that L2 learners in EFL contexts are mostly exposed to written text modality rather than the spoken text modality. Therefore, supporting this input mode with other textual enhancements, such as listening while presenting the text in writing and listening with audiovisual input, may lead to better processing and learning gain.

It is, therefore, clear that studies that have explored vocabulary learning from listening are scarce, especially in the Saudi EFL context. Moreover, no previous study has used aural tests to gauge vocabulary gains from spoken input, e.g., listening only. Thus, this poses a research gap in instrumentations that the current study tries to fill.

Incidental learning of vocabulary through reading while listening

The interest in exploring the effect of reading while listening on incidental vocabulary learning has been noted in recent literature (e.g., Brown et al., 2008; Van Vu & Peters, 2020; 2021; Webb & Chang, 2022). The study by Brown et al. (2008), reported earlier, was one of the earliest studies to examine the impact of reading while listening on vocabulary gains. In a later study, Webb and Chang (2022) further examined vocabulary learning from reading only and reading while listening with a group of Taiwanese EFL learners. The study results indicated that the learners gained more vocabulary through reading while listening than in the reading-only condition.

For a more general view of developing L2 skills, Chang and Millett (2013) looked at developing L2 listening fluency through extensive listening to audio recordings from graded readers in three conditions: reading only, reading while listening, and listening only. The results have revealed that reading while listening resulted in significantly better listening fluency than the reading-only and listening-only modes.

The literature suggests that dual modes input, such as reading while listening, is more advantageous to L2 learners than single modes, i.e., listening only or reading only. However, results remain mixed, and further explorations are needed to advance our knowledge on this matter. One prominent methodological limitation in most previous studies is using written test measures to examine incidental vocabulary gains from aural input. Therefore, instruments that mainly measure aural modality could be operationalized to sensibly quantify any potential vocabulary gain from aural input. This is believed to be an essential methodological shift, and the current study tries to fill this methodological gap.

Incidental learning of vocabulary through listening with visual enhancement

Dual Code Theory (DCT) assumes that "mental representations are associated with theoretically distinct verbal and nonverbal symbolic modes and retain properties of the concrete sensorimotor events on which they are based" (Clark & Paivio, 1991: 151). The verbal code concerns language, while the nonverbal code concerns mental imagery. From a vocabulary perspective, it is believed that if learners are exposed to vocabulary through DCT, their cognitive load is reduced, making more space in working memory capacity, which in turn would lead to better learning. Thus, many studies believe that exposing words to learners through both verbal and nonverbal channels are better learned and retained (e.g., Clark & Paivio, 1991; Sadoski, 2005; and Mayer, 2014). According to Clark and Paivio (1991: 151-152), DCT verbal system covers visual, auditory, articulatory, and other modality-specific verbal codes, while nonverbal representations include modality-specific images for shapes, environmental sounds, actions, skeletal or visceral sensations related to emotion, and other nonlinguistic objects and events.

The combination of visual and auditory modalities has recently attracted researchers' attention in the L2 vocabulary learning endeavor since it seems that this combination could lead to more and better learning of vocabulary. For instance, the literature points out some evidence for the value of watching videos or TV in developing vocabulary knowledge in L2 learners. However, most of the research in this area has not directly focused on visual enhancement (i.e., colors, pictures, and highlighting), but instead on vocabulary learning from audiovisual input enhanced with different types of subtitling (L1 or L2 subtitles) (e.g., Peters, Heynen, and Puim`ege, 2016; Winke, Gass, and Sydorenko, 2010). An exception is the two studies by Peters and Webb (2018) and Rodgers and Webb (2020). For example, Rodgers and Webb (2020) examined the

potential for incidental vocabulary gain from several TV series episodes. Their results indicated that six words were learned after watching the TV series as measured through pretest and posttest methods.

The study of Peters and Webb (2018) examined incidental vocabulary learning through watching a full-length, one-hour documentary. Similar to Rodgers and Webb (2020), Peters and Webb's study employed a pretest and posttest design. Vocabulary learning was measured through the target items' aural-form matching and meaning recall. The results showed that approximately four words were learned after viewing the documentary. Results from Rodgers and Webb's (2020) and Peters and Webb's (2018) studies are in line with those from previous studies, such as Perez, Peters, Clarebout, and Desmet (2014) and Sydorenko (2010), which have looked at incidental vocabulary learning from short-length videos enhanced with subtitles.

From the preceding discussion, it is evident that the literature has mixed results regarding the amount of gains an input mode can cater to incidental vocabulary learning. Therefore, more research is needed to investigate this area, especially the reading while listening and listening only modes. More importantly, all previous studies have not employed aural test measurements to gauge participants' gains of new items. Therefore, our study is to fill those gaps to better understand the effects of different input modes on incidental learning by answering the following research questions (RQs):

> **RQ1.** To what extent does incidental vocabulary learning occur through different input modes: Listening only, reading while listening, and listening with visual enhancement?
> **RQ2.** How does vocabulary gain compare across meaning recognition, aural-form matching, and meaning recall? Is there an observed order of vocabulary learning at these levels of word knowledge?

Methodology

Participants

The participants were initially 135 Saudi EFL learners from a public intermediate (middle) school in Riyadh, Saudi Arabia, ranging between 15 and 16 years old. They were all male pupils in their 9th grade (i.e., the third intermediate year of the Saudi Arabian education system). They had four English sessions per week, with 45 minutes allotted for each session. However, due to the Covid-19 pandemic, time per session was reduced to 35 minutes. The participants had studied English subjects in their formal education for about five years. However, since some participants were already familiar with one or more of the target words, 45 participants were removed from the study at the screening stage. Therefore, the number of participants included in the study was reduced to 90. Furthermore, due to the restrictions set by the Saudi Ministry of Education during Covid-19 pandemic, not many school principals were willing to conduct any research study since schooling time was very restricted. Therefore, conducting the study on these 90 participants was not so convenient (see the limitations below for more details).

The book they were using to study English was Full Blast 6 (a textbook designed specifically for the Saudi Ministry of Education by MM Publications). As reported by MM Publications, Full Blast is a B2 level according to the Common European Framework of Reference for Languages (CEFR). However, the language teacher stated that most of his participating learners had weak English language proficiency. Therefore, we administered the first 1000-word level of Webb, Sasao, and Ballance (2017) updated Vocabulary Levels Test to the participants, and their level, on average, appeared near the mastery of the 1000-level but fully

mastered that level (*M* = 25.27, *SD* = 12.31). So, they are, in fact, at a low proficiency level.

Materials

Based on the participants' scores on Webb et al.'s (2017) Vocabulary Levels Test (VLT) (1,000 word-level) and the teacher's recommendation concerning the students' low level of English, the target words were chosen from a non-fiction monologue A1-level graded reader to be appropriate for the participants' level. The book was one of the Oxford Read and Discovered series accompanied by an audio enhancement for reading the text. The speech speed was about 80 words per minute. Therefore, the speech speed was suitable for A1-level learners. The chosen title was *Camouflage*. It was believed that this topic could be new to our participants; therefore, the content would not be familiar to them.

Furthermore, it was speculated that a fact file reader (i.e., non-fiction) would be easier for low-level learners to follow than any narrative with multiple characters and divergent plots. As discussed in the procedures section later, all the experimental groups had the same content but different modes. According to the book's blurb, Camouflage has a word count of 843 words. Furthermore, according to our book analysis on VocabProfiler (https://www.lextutor.ca/vp/), the book has 162-word families, and around 85% of the tokens (i.e., 122-word families) fall within the first 1000-words band.

Selection of target items

As mentioned before, the words were given to the participants for screening purposes. Since they were familiar with many of the target words, some participants were removed from the study at the first screening stage. This screening resulted in removing 45 participants from the study. Hence, 16 words were used as target words for the study. They included 13 nouns and three verbs (see Table 1 for the list of the target words, their types, levels according to the BNC and

COCA, and frequency of occurrence). These 16 target words were used for the design of the study measures. The frequency of occurrence of the 16 target items is presented in Table 1. Since the study primarily investigated vocabulary learning uptake from a listening mode, the final list of the 16 target words was voice-recorded with 8-second intervals between each of them. This allowed the participants to perform the required tasks after hearing each word. The words were introduced to the participants in a spoken format

Table 1.

Target Words, their Type, their Level, and their Frequency of Occurrence in the Text

Target word	Type	Word level according to BNC/COCA	Frequency of occurrence
feather	N	2	7
coral	N	5	3
pattern	N	2	2
gecko	N	14	1
fur	N	2	6
stripe	N	4	3
leaf	N	4	12
seal	N	2	2
bottom	N	1	1
shape	N	1	1
discover	V	1	5
plant	N	1	3
melt	V	2	1
gazelle	N	12	7
beetle	N	6	5
bounce	V	2	1

Instruments

Tests used in the current study are believed to gauge aural vocabulary gains. Previous studies on learning vocabulary from aural input have used tests that assess the orthographic knowledge of test-takers before and after the treatment (e.g., Brown et al., 2008; Feng & Webb, 2020; van Zeeland & Schmitt, 2013; Vidal, 2011). It is argued that using orthographic-based tests to measure learners' vocabulary uptake from aural input could be insensitive to eliciting vocabulary gains. In the current study, efforts have been made to design tests geared toward measuring aural incidental vocabulary learning. Therefore, all the tests used in the study were presented to the participants in spoken formats.

Meaning-recall test

To minimize the potential learning effect of testing as much as possible, we administered the meaning recall test only as a pre-test. This test format was also used as a posttest to quantify any potential vocabulary gain. In this meaning recall test, the test-takers were first presented with a recording and asked to decide whether they knew the target word (Yes, I know it) or (No, I do not know it). In the second step, if the test-taker responded Yes to the given item, the learner must translate it into Arabic (meaning recall). A target word was marked one if correctly translated into Arabic or 0 if not. Since the participants were not familiar with the target words before the intervention, they were only expected to provide the specific meaning of polysemous target words as intended in the text. That is, if a student provided a correct meaning that did not fit the context of the passage, a zero mark was given.

Furthermore, since the words in the text were carefully controlled, all of the words appeared with the same meaning throughout the whole text. However, when scoring, spelling mistakes and parts of speech were overlooked as long as the answer conveyed the intended meaning. The test was scored by two raters who are native-Arabic

speakers and professors of English. The inter-rater reliability rate was measured simply. If the two raters agreed on an item, the item was given 1; if not, the item was given 0. The scores of all items were then accumulated. No disagreement was observed between the two raters.

Example item:

[Feather]:

No, I do not know it [ ].

Yes, I know it [ ].

If Yes, give an Arabic translation [……]. Then, the correct answer is [ريشة].

Meaning-recognition test

This test was a four-option multiple-choice test. The target word was read to the participants (the same previous recording) for each item, and they were provided with four L1 (Arabic) words to choose from. All the other three distractors had the same part of speech, singular/plural form, or tense as the target word.

Example item:

[Feather]:

a. ريشة [feather]
b. أب [father]
c. نبتة [plant]
d. وزن [weight]

Aural-form matching test

In this test format, the participants were shown an L1 word, and they heard three English words (including the target word) to choose from. The other two distractors were chosen to sound similar to the target word regardless of their parts of speech. The participants were required to choose a, b, or c for each prompt word. No orthographic representations of the options were provided. The whole test was

delivered in the spoken format except for the Arabic words.

Although the test only had three options, we aimed to use phonologically similar words to the target word. However, not all target words share significant phonological similarities with many words. Therefore, we deemed two distractors sufficient, acknowledging that this could make the test challenging. Having just three options could be a limitation that future research could consider.

Example item:

[ريشة]:

a. father
b. further
c. feather [correct answer]

The study procedure

The study's first phase included administering the meaning-recall test as a pretest and the 1000-word level of the updated VLT (Webb et al., 2017). Then, following the screening criteria of the target items and the participants described earlier in the paper, the participants were divided into three experimental groups (30 participants in each group): Listening Only, reading-while-listening, and listening with visual enhancements. The listening-Only group listened to the book's audio file, i.e., without reading/seeing any written text. The reading-while-listening group had the same audio content but with written L2 text available on a screen. The text was a Times New Roman, 14-pt size. No colors or highlights were added to the text. There was only a black background to make the text readable for the participants. The listening with visual enhancements group had the original book content with various enhancements (e.g., colors, pictures, and some highlighted words; made initially by the publisher). The audio file was played simultaneously as the book pages were browsed through. The book (i.e., Camouflage) is suitable for beginning readers. The font size is larger than the average book font size; every page is divided into two halves. One of the two

halves has just around seven lines of about 6-8 words each, while the other half has one background picture or more. Each of these pictures has a highlighted, captioned word of the name of an animal or a short phrase describing an essential element in the picture. So, the enhancements meant in the current study are the colorful pages, large fonts, and captioned pictures (no target words were in these captioned pictures).

After the treatment was completed, the three posttests described earlier were immediately administered to the participants. The tests were administered to the participants in the following order, (1) meaning-recall test, (2) meaning-recognition test, and (3) aural-form matching test, in an attempt to reduce the testing effect as much as possible. All the three measures were presented to the participants in spoken form electronically using Google Forms. Participants' informed consent was obtained through a statement in the Forms, ensuring that all provided information and answers would remain confidential and be used solely for research purposes. In addition, all the instructions in posttests were introduced in Arabic to ensure that the participants fully understood the required tasks.

The ultimate aim of the study was to find answers to the following research questions:

1. To what extent does incidental vocabulary learning occur through different input modes: Listening only, reading while listening, and listening with visual enhancement?
2. How does vocabulary gain compare across meaning recognition, aural-form matching, and meaning recall? Is there an observed order of vocabulary learning at these levels of word knowledge?

Results

The findings are presented according to the two research questions as follows:

Preliminary analysis

All data were analyzed using SPSS (V. 28). Descriptive statistics of scores from the three groups on meaning recall, meaning recognition, and aural-form matching measures are presented in Table 2. Good reliability for the measures (α = .85, .87, .84, respectively), as assessed with Cronbach's alpha, was observed. Table 2 showed evidence of incidental vocabulary learning from the three input modes (listening only, reading while listening, and listening with visual enhancement) at the three levels of tested word knowledge. The target words were unknown to the participants prior to the intervention, as as the pre-test results indicated. Therefore, any reported vocabulary gain is considered a development from no knowledge of the target words on the pre-tests to some knowledge on the post-tests.

Overall, the results showed increased mean scores on meaning recognition, aural-form matching, and meaning recall from the three input modes. Mean score differences were observed both between and within the groups across the three types of tested word knowledge.. Inferential analyses were then conducted to examine whether the observed differences were statistically significant. SPSS (V. 28) was employed for data analysis, with a 95% Confidence Interval applied to all analyses.

Table 2.

Descriptive Statistics of Posttest Scores by Input Mode

Input mode	post-tests: *M* (*SD*)*		
	Meaning recognition	Aural-Form recognition	*Meaning recall*
L	6.60 (2.31)	4.83 (1.87)	1.47 (.81)
RL	8.60 (3.57)	6.90 (2.80)	2.33 (.64)
LV	12.00 (4.05)	9.83 (3.98)	5.07 (1.58)

Note. L = Listening only condition; RL = Reading while listening condition; LV = Listening with visual enhancement condition. The number of participants is 30 in each input condition. * maximum = 16.

RQ1: To what extent does incidental vocabulary learning occur through different input modes: Listening only, reading while listening, and listening with visual enhancement?

Descriptive statistics reported in Table 2 generally show a positive effect of the three input modes on incidental vocabulary learning (mean scores are illustrated in Figure 1). However, a relatively greater incidental vocabulary gain from the listening with visual enhancement condition was observed at the levels of meaning recognition, aural-form matching, and meaning recall. To examine whether the differences between the groups were statistically significant, an analysis of variance (ANOVA) was performed. Results of a one-way between-groups ANOVA indicated an overall significant difference in the meaning recognition, aural-form matching, and meaning recall scores, $F(2, 87) = 19.45$, $p < .001$, $\eta_p^2 = .31$; $F(2, 87) = 12.04$, $p < .001$, $\eta_p^2 = .22$; $F(2, 87) = 7.53$, $p = .001$, $\eta_p^2 = .02$, respectively.

Post hoc comparisons using Bonferroni correction showed that learners in the listening with visual enhancement group had

significantly higher scores on the meaning recognition test than those reading while listening and listening only groups (p = .001, d = .89; p < .001, d = 1.64, respectively). However, the difference between reading while listening and listening only groups was not statistically significant. Similarly, the results showed that the listening with visual enhancement group performed significantly better in the aural-form matching test than the reading while listening and listening only groups (p = .02, d = .73; p < .001, d = 1.44, respectively). Likewise, although the results indicated that learners in the reading while listening group achieved better scores in the aural-form matching test, the difference was not statistically significant. Additionally, similar results were revealed for the meaning recall condition. Again, the listening with visual enhancement group significantly outperformed the listening only and reading while listening groups (p = .02, d = .67; p = .001, d = .95, respectively). No significant difference was found between listening-only and reading-while-listening groups.

These results suggest that listening with visual enhancement is a valuable mode of input to support incidental vocabulary learning, over and above listening only and/or reading-while-listening. Nonetheless, close examination of the results interestingly reveals that listening only and reading while listening modes of input also contributed to vocabulary gain, at least in terms of meaning and aural-form matching levels.

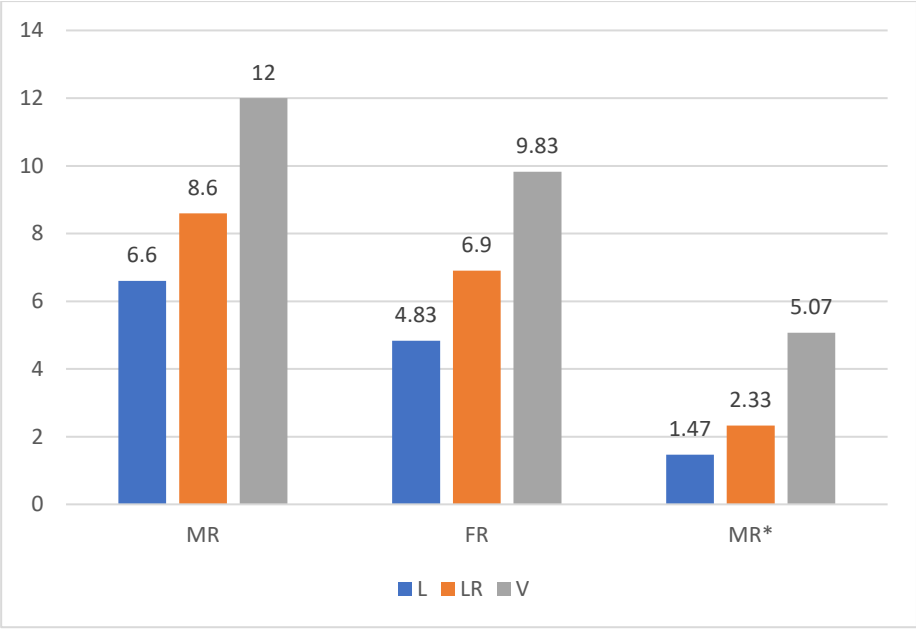

Figure 1. *Incidental Vocabulary Gains Across Input Modes at the Different Levels of Word Knowledge*

Note. MR = meaning recognition; FR = aural-form matching; MR* = meaning recall.

RQ2: How does vocabulary gain compare across meaning recognition, aural-form matching, and meaning recall? Is there an observed order of vocabulary learning at these levels of word knowledge?

To answer the second research question, a paired samples *t*-test was carried out to examine how each input mode contributes to the learning of vocabulary at the three levels of word knowledge (i.e., meaning recognition, aural-form matching, and meaning recall). Interestingly, the results revealed statistically significant differences in scores for meaning recognition, aural-form matching, and meaning recall across the three conditions. First, a statistically

significant difference was found between scores on the meaning recognition and aural-form matching (pair 1) within each group: listening only condition, $t(29) = 3.14$, $p < .01$; reading while listening, $t(29) = 2.79$, $p < .01$; listening with visual enhancement, $t(29) = 3.07$, $p < .01$. Second, a significant difference was also found between meaning recognition and meaning recall (pair 2): listening only, $t(29) = 8.39$, $p < .001$; reading while listening, $t(29) = 11.38$, $p < .001$; listening with visual enhancement, $t(29) = 9.02$, $p < .001$. Finally, the difference between aural-form matching and meaning recall was also found statistically significant: listening only, $t(29) = 7.30$, $p < .001$; reading while listening, $t(29) = 7.54$, $p < .001$; listening with visual enhancement, $t(29) = 6.72$, $p < .001$. These results suggest that although the three aural input modes facilitated vocabulary gain, their contributions to the levels of word knowledge differed. The results from the pairwise comparisons (further depicted in Figure 2) indicated that, in each input mode, the learners showed a better gain of vocabulary at the meaning recognition level, followed by aural-form matching and meaning recall. Interestingly, the results illustrated in Figure 2 showed that listening with visual enhancement led to comparatively greater gains across the three levels of word knowledge than the other two input modes. Furthermore, the results showed an order of incidental learning of vocabulary where learners appear to gain vocabulary at the level of meaning recognition, aural-form matching, and then meaning recall. The meaning recall seems to be acquired at a later stage of vocabulary development.

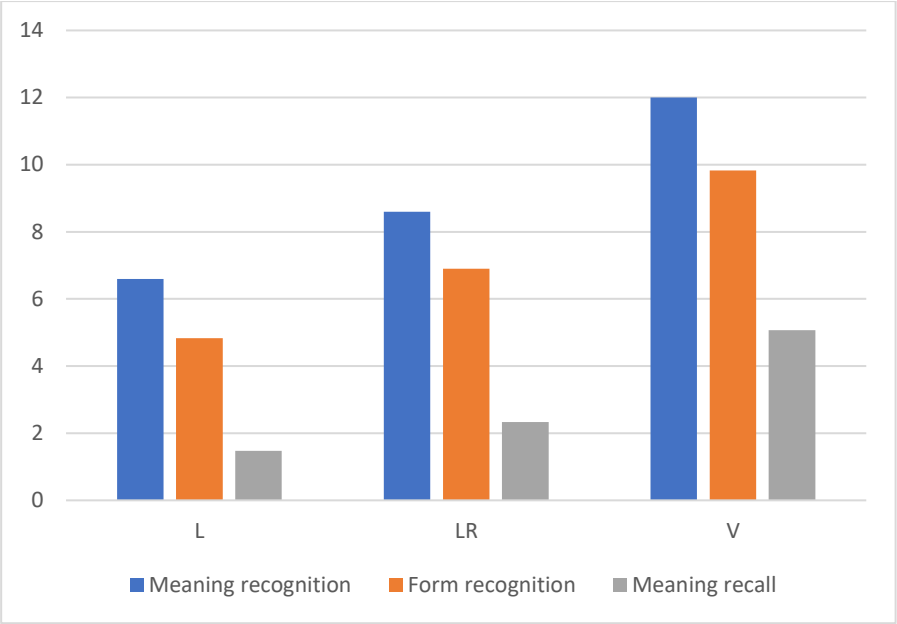

Figure 2. *Incidental Vocabulary Gains Across the Different Levels of Word Knowledge*

Discussion

The present study examined the incidental vocabulary learning through three modes of aural input: listening only, reading while listening, and listening with visual enhancement. It also explored the potential order of vocabulary learning at the levels of meaning recognition, aural-form matching, and meaning recall.The study is in line with the recent calls (e.g., Schmitt, 2019) to explore the impact audiovisual input has on vocabulary learning and instrumental enhancements that facilitate incidental vocabulary learning. To appropriately quantify incidental vocabulary learning from aural-oriented input mode, we employed relevant measures that tap into aural modality.

This study helps to address research gaps in two ways. First, to the researchers' knowledge, the study is the first to compare incidental vocabulary learning utilizing aural-written, audio, and

audiovisual input. Also, unlike most previous studies exploring this matter, the study operationalized measures that elicit responses in *aural* formats to match the task modality.

In answer to the first research question, the findings showed that notable vocabulary learning did occur from reading while listening, listening only, and viewing with L2 subtitling. The results support findings from earlier research, showing that vocabulary was incidentally gained via reading while listening (e.g., Valentini, Ricketts, Pye, and Houston-Price, 2018), listening only (e.g., Feng & Webb, 2020), and viewing (e.g., Feng & Webb, 2020; Perez, 2020). The findings also indicated that the three input modes contributed to vocabulary learning at the meaning recognition, aural-form matching, and meaning recall. However, while vocabulary was evident through three different input modes, listening with visuals resulted in the most significant gain over the three levels of word knowledge. This outcome may support the argument that L2 learners prefer viewing in the target language. In addition, this mode of input increases their motivation to learn new vocabulary (e.g., Gieve & Clark, 2015; Rodgers, 2013).

Reinforcing the beneficial impact of well-studied input modes, (i.e., reading while listening and listening) on incidental vocabulary learning, the results of viewing graded non-fiction readers appear particularly useful, as the foci of several earlier studies have been on specialized materials (e.g., Vidal, 2003, 2011), and short videos (e.g., Perez et al., 2014; Sydorenko, 2010), which are, arguably, a less common form of viewing with L2 learners. It is particularly interesting that this research indicates viewing L2 graded readers contributes to vocabulary gain because this form of viewing is believed appropriate for L2 learners where the materials can be adapted to cater to their proficiency level.

Additionally, it is most likely that the more enhancements a text has, the more gains of vocabulary knowledge could occur. In the current study, the participants in the viewing mode group obtained

the best gains as extra enhancements were present, i.e., pictures, texts, colors, different font sizes, and paragraph-type text organization. Learning is highly affected by activating different senses and skills simultaneously. Besides, more enhancements added to the text/words (e.g., bold-faced, italics, colors, and pictures) are believed to lead to deeper text processing. This is in line with Nation's (2015, p. 136) statement that the "deeper the quality of the meetings, the more likely learning is to occur ."Nation (2001, p. 199) lists using pictures as one of the features of deep processing that he describes as "very good ."Our study and others report better results for participating groups exposed to more text enhancements (e.g., Feng & Webb, 2020; Peters & Webb, 2018; Rodgers & Webb, 2020; Yeldham, 2018). This goes in hand with what Multimedia Learning Theory and Dual Code Theory suggest, “people learn more deeply from words and pictures than from words alone” (Mayer, 2014, p. 47).

The results of the second research question revealed a consistent pattern of vocabulary learning at the three levels of word knowledge (meaning recognition, aural-form matching, and meaning recall) across input modes. In each mode, a significant difference was observed between the levels of word knowledge, where meaning recognition showed the most significant gain and meaning recall the least. This supports van Zeeland and Schmitt (2013) in that the participants developed an order of vocabulary knowledge that moves from meaning recognition into meaning recall. However, the comparison between groups indicated that the viewing condition has resulted in the most significant vocabulary gain across the meaning recognition, aural-form matching, and meaning recall. However, the difference between reading while listening and listening only conditions were insignificant. The difference between reading while listening and listening only modes of input appear surprising and contrasts that of Valentini et al. (2018). In their study, Valentini et al. found that the combined aural and written modalities led to better

vocabulary development than aural input alone. One possible interpretation of the result from the current study is that the participants were at a relatively low level. Hence, their ability to decode the phonological form is still limited, benefiting less from the reading enhancement of the text.

Interestingly, our results suggest a higher incidental vocabulary learning from viewing than in previous studies (e.g., Perez, 2020; Peters & Webb, 2018). In this study, an absolute gain of 75% of the target words was found in the meaning recognition, which is substantially higher than that found in Perez (2020), about 20% of the target words, and Peters and Webb (2018), about 14%. In terms of the meaning recall, the results of our study revealed a gain of about 32% of the target words, which is much greater than that found in Perez (2020) and Peters and Webb (2018), 2.1%, 8%, respectively. Similar to the studies mentioned above, meaning recall turned out to be the most challenging level of word knowledge. However, this is not surprising, as previous research has indicated that numerous encounters with words are needed before L2 learners can recall a word meaning from incidental exposure to listening (Brown et al., 2008).

The more considerable gain of vocabulary found in our study than that in Perez (2020) and Peters and Webb (2018), for example, might be attributed to the fact that our target items were from the A1 graded reader level. It is most likely that words at this level are relatively easier to learn than those from a documentary (Peters & Webb, 2018) and pseudowords (Perez, 2020). A second explanation for the more considerable vocabulary learning in this study is probably related to the speech rate. A slower speech rate is usually used in materials delivered to A1-level learners. This might have helped the learners in this study to uptake the target items. Thirdly, and most importantly, when appropriate measures that tap the modality of the target words are used, at least partial knowledge can be rightly captured. Using written measures to elicit aural uptake

may not well pick up a low level of knowledge acquired from little exposure to the language in an L2 context.

Pedagogical implications

The findings of this study suggest a notable gain of vocabulary occurring incidentally from listening only, reading while listening, and viewing. Rather than focusing on only one mode of input in the language classroom, it would be helpful for language teachers to acquaint their learners with various modes of input to support their learning and comprehension of the materials. Employing such a practice would raise the learners' awareness that language learning is a matter of relying on multiple types of input. As the classroom is time-constrained, and large amounts of input are necessary for some words to be learned through meaning-focused input, developing the learners' skills to apply different input modes to support their learning outside of the language classroom would be invaluable to their language development. Additionally, although much research supports reading as a key contributor to vocabulary development, this study and others, such as Feng & Webb (2020), provide empirical evidence that listening, reading while listening, and viewing also positively impact lexical development.

Conclusion

This study offered evidence that listening, reading while listening, and viewing could contribute significantly to incidental vocabulary learning from graded readers. The study's findings suggest no statistically significant difference between listening and reading-while listening input modes, though slightly better vocabulary was observed in the reading-while-listening mode. The viewing, however, resulted in a notably more significant vocabulary gain than the other two conditions. Interestingly, the results suggest that vocabulary learning occurred at the meaning recognition, aural-form matching, and meaning recall of word knowledge from the three input modes but was relatively greater from the viewing input.

An order of vocabulary learning was also observed in this study, developing from meaning recognition to meaning recall. While reading programs to promote vocabulary learning has long been advocated, it is recommended that analogous listening/viewing programs are also developed.

However, reading the results of the current study could take into account that the target words were all of a concrete nature, which in turn could have led to our learners' higher vocabulary learning rates than previous studies. For instance, Hargis and Gickling (1978) found that participating children learned concrete words 12% faster than abstract words. Word parts are also believed to play a role in vocabulary learning (e.g., Chang, 2019; Ellis and Beaton, 1993; and Peters, 2020). For example, Ellis and Beaton (1993) state that a word part of speech affects its learnability and that "nouns are the easiest to learn" (p. 565).

There are several limitations in this study that need to be addressed. First, the number of the target words was relatively small due to the elimination of 19 words, which is a considerable loss. Such many words could have led to more excellent uptake rates in our study. Therefore, future research could benefit from including more target words to more thoroughly explore uptake rates. Second, our initial intention was to run some frequency analysis of the 37 target words, but since we lost 19 words, we ended up with irregular intervals of frequency bands in the finalized list of the 16 words. Therefore, we suggest that future research could aim at exploring this area as previous research has shown that frequency affects learners' vocabulary learning (e.g., Fernández & Schmitt, 2015; Nation, 2001; Schmitt, 2000)

Furthermore, the aim was to rerun the three posttests after two weeks to examine vocabulary learning retention. However, to our surprise, despite being helpful, the teacher apologized for not being able to conduct more tests due to restricted class time during the Covid-19 pandemic and strict Ministry of Education

regulations. Pupils had to finish the textbook assigned by the Ministry and that the posttests would consume a great deal of their class time. Hence, we could not explore retention gains from aural input. Nevertheless, we believe that running delayed posttests would inform our understanding of the vocabulary gain retention from aural input, which needs to be explored in future research.

Additionally, we should have introduced the meaning recognition, not the meaning recall, test as a pre-test as the former is less demanding than the latter, and that some learners could have been more able to recognize some aspects of the target words than being able to recall their meanings.

Conflict of interest

The author declares that he has no conflict of interest.

Informed consent

Informed consent was obtained from all individual participants included in the study.

Funding

The author did not receive support from any organization for the submitted work.

References

Akhtar, N. (2004). Contexts of early word learning. In Hall, D. & Waxman, S. R. (Eds.), *Weaving a Lexicon* (pp. 485–507). MIT Press.

Alderson, J. C. (2005). *Diagnosing foreign language proficiency: The interface between learning and assessment*. A&C Black.

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

Clark, J., & Paivio, A. (1991). Dual coding theory and education. *Educational Psychology Review*, *3*, 149 – 210.

Chang, A., & Millett, S. (2013). The effect of extensive listening on developing L2 listening fluency: Some hard evidence. *ELT Journal*, *68*(1), 31-40.

Chen, Y. (2021). Comparing incidental vocabulary learning from reading-only and reading-while-listening. *System*, *97*, 1-13.

Feng, Y., & Webb, S. (2020). Learning vocabulary through reading, listening, and viewing: Which mode of input is most effective? *Studies in Second Language Acquisition*, *42*(3), 499-523.

Gieve, S., & Clark, R. (2005). The Chinese approach to learning: Cultural trait or situated response? The case of a self-directed learning programme. *System*, *33*(2), 261-276. http://doi.org/10.1016/j.system.2004.09.015.

Henderson, L., Devine, K., Weighall, A., & Gaskell, G. (2015). When the daffodat flew to the intergalactic zoo: Off-line consolidation is critical for word learning from stories. *Developmental Psychology*, *51*(3), 406-417.

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

Laufer, B. (2005). Focus on form in second language vocabulary learning. *Eurosla Yearbook*, *5*(1), 223–250.

Mayer, R. E. (2014). Cognitive theory of multimedia learning. In R. E. Mayer (Ed.), *The Cambridge handbook of multimedia learning* (2 ed., pp. 43-71). Cambridge University Press.

Perez, M. (2020). Incidental vocabulary learning through viewing video: The role of vocabulary knowledge and working memory. *Studies in Second Language Acquisition*, *42*(4), 749-773.

Perez, M., Peters, E., Clarebout, G., & Desmet, P. (2014). Effects of captioning on video comprehension and incidental vocabulary learning. *Language Learning & Technology*, *18*, 118-141.

Nation, I. S. P. (2001). *Learning vocabulary in another language*. Cambridge University Press.

Nation, I. S. P. (2015). Principles guiding vocabulary learning through extensive reading. *Reading in a Foreign Language*, *27*, 136-145.

Nation, I. S. P., & Waring, R. (2020). *Teaching extensive reading in another language*. Routledge

Neuman, S. B., & Koskinen, P. (1992). Captioned television as "comprehensible input": Effects of incidental word learning from context for language minority students. *Reading Research Quarterly*, *27*, 95-106.

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

Peters, E. (2018). The effect of out-of-class exposure to English language media on learners' vocabulary knowledge. *ITL—International Journal of Applied Linguistics*, *169(1)*, *142*-168.

Peters, E., & Webb, S. (2018). Incidental vocabulary acquisition through viewing L2 television and factors that affect learning. *Studies in Second Language Acquisition*, *40*(3), 551–577.

Peters, E., Heynen, E., & Puim`ege, E. (2016). Learning vocabulary through audiovisual input: The differential effect of L1 subtitles and captions. *System*, *63*, 134-148.

Renandya, W. A., & Farrell, T. S. (2011). "Teacher, the tape is too fast!" Extensive listening in ELT. *ELT Journal*, *65*, 52–59.

Rodgers, M. P. H. (2013). *English language learning through viewing television: An investigation of comprehension, incidental vocabulary acquisition, lexical coverage, attitudes, and captions* [Unpublished doctoral dissertation]. Retrieved from http://researcharchive.vuw.ac.nz/xmlui/bitstream/handle/10063/2870/thesis.pdf?sequence52.

Rodgers, M., & Webb, S. (2020). Incidental vocabulary learning through viewing television. *ITL - International Journal of Applied Linguistics*, *171*(2), 191-220.

Sadoski, M. (2005). A Dual Coding View of Vocabulary Learning. Reading & Writing Quarterly: Overcoming Learning Difficulties, *21*(3), 221-238.

Schmitt, N. (2019). Understanding vocabulary acquisition, instruction, and assessment: A research agenda. *Language Teaching, 52(2),* 261-274.

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

Valentini, A., Ricketts, J., Pye, R., & Houston-Price, C. (2018). Reading while listening promotes word learning from stories. *Journal of Experimental Child Psychology, 167,* 10-31.

van Zeeland, H., & Schmitt, N. (2013). Incidental vocabulary acquisition through L2 listening: A dimensions approach. *System*, *41*, 609-624.

Vidal, K. (2003). Academic listening: A source of vocabulary acquisition? *Applied Linguistics*, *24(1)*, 56-89.

Vidal, K. (2011). A comparison of the effects of reading and listening on incidental vocabulary acquisition. *Language Learning*, *61(1)*, 219-258.

Van Vu, D., & Peters, E. (2020). Learning vocabulary from reading-only, reading-while-listening, and reading with textual input enhancement: Insights from Vietnamese EFL learners. RELC journal, 53(1), 85-100.

Webb, S., & Chang, A. (2022). How does mode of input affect the incidental learning of collocations? *Studies in Second Language Acquisition*, *44*(1), 35-56.

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

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*(1), 33-69.

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

Yeldham, M. (2018). Viewing L2 captioned videos: What's in it for the listener? *Computer Assisted Language Learning*, *31*(4), 367-389.

Zhang, P., & Graham, S. (2020). Learning Vocabulary Through Listening: The Role of Vocabulary Knowledge and Listening Proficiency. *Language Learning*, *70*(4), 1017-1053.