

# Incorporating technology for teaching vocabulary to second language learners

**Edited by**

Ehsan Rassaei, Xuesong Gao and Brian Poole

**Published in**

Frontiers in Psychology



## FRONTIERS EBOOK COPYRIGHT STATEMENT

The copyright in the text of individual articles in this ebook is the property of their respective authors or their respective institutions or funders. The copyright in graphics and images within each article may be subject to copyright of other parties. In both cases this is subject to a license granted to Frontiers.

The compilation of articles constituting this ebook is the property of Frontiers.

Each article within this ebook, and the ebook itself, are published under the most recent version of the Creative Commons CC-BY licence. The version current at the date of publication of this ebook is CC-BY 4.0. If the CC-BY licence is updated, the licence granted by Frontiers is automatically updated to the new version.

When exercising any right under the CC-BY licence, Frontiers must be attributed as the original publisher of the article or ebook, as applicable.

Authors have the responsibility of ensuring that any graphics or other materials which are the property of others may be included in the CC-BY licence, but this should be checked before relying on the CC-BY licence to reproduce those materials. Any copyright notices relating to those materials must be complied with.

Copyright and source acknowledgement notices may not be removed and must be displayed in any copy, derivative work or partial copy which includes the elements in question.

All copyright, and all rights therein, are protected by national and international copyright laws. The above represents a summary only. For further information please read Frontiers' Conditions for Website Use and Copyright Statement, and the applicable CC-BY licence.

ISSN 1664-8714  
ISBN 978-2-83251-977-6  
DOI 10.3389/978-2-83251-977-6

## About Frontiers

Frontiers is more than just an open access publisher of scholarly articles: it is a pioneering approach to the world of academia, radically improving the way scholarly research is managed. The grand vision of Frontiers is a world where all people have an equal opportunity to seek, share and generate knowledge. Frontiers provides immediate and permanent online open access to all its publications, but this alone is not enough to realize our grand goals.

## Frontiers journal series

The Frontiers journal series is a multi-tier and interdisciplinary set of open-access, online journals, promising a paradigm shift from the current review, selection and dissemination processes in academic publishing. All Frontiers journals are driven by researchers for researchers; therefore, they constitute a service to the scholarly community. At the same time, the *Frontiers journal series* operates on a revolutionary invention, the tiered publishing system, initially addressing specific communities of scholars, and gradually climbing up to broader public understanding, thus serving the interests of the lay society, too.

## Dedication to quality

Each Frontiers article is a landmark of the highest quality, thanks to genuinely collaborative interactions between authors and review editors, who include some of the world's best academicians. Research must be certified by peers before entering a stream of knowledge that may eventually reach the public - and shape society; therefore, Frontiers only applies the most rigorous and unbiased reviews. Frontiers revolutionizes research publishing by freely delivering the most outstanding research, evaluated with no bias from both the academic and social point of view. By applying the most advanced information technologies, Frontiers is catapulting scholarly publishing into a new generation.

## What are Frontiers Research Topics?

Frontiers Research Topics are very popular trademarks of the *Frontiers journals series*: they are collections of at least ten articles, all centered on a particular subject. With their unique mix of varied contributions from Original Research to Review Articles, Frontiers Research Topics unify the most influential researchers, the latest key findings and historical advances in a hot research area.

Find out more on how to host your own Frontiers Research Topic or contribute to one as an author by contacting the Frontiers editorial office: [frontiersin.org/about/contact](https://frontiersin.org/about/contact)

# Incorporating technology for teaching vocabulary to second language learners

## Topic editors

Ehsan Rassaei — Majan University College, Oman

Xuesong Gao — University of New South Wales, Australia

Brian Poole — Majan University College, Oman

## Citation

Rassaei, E., Gao, X., Poole, B., eds. (2023). *Incorporating technology for teaching vocabulary to second language learners*. Lausanne: Frontiers Media SA.  
doi: 10.3389/978-2-83251-977-6

# Table of contents

04	<b>Editorial: Incorporating technology for teaching vocabulary to second language learners</b> Ehsan Rassaei
07	<b>Exploring the Effect of Assisted Repeated Reading on Incidental Vocabulary Learning and Vocabulary Learning Self-Efficacy in an EFL Context</b> Habib Soleimani, Farnoosh Mohammaddokht and Jalil Fathi
18	<b>Vocabulary Mobile Learning Application in Blended English Language Learning</b> Petra Polakova and Blanka Klimova
28	<b>On-Screen Texts in Audiovisual Input for L2 Vocabulary Learning: A Review</b> Rong Wei and Lin Fan
39	<b>Teaching Academic Words With Digital Flashcards: Investigating the Effectiveness of Mobile-Assisted Vocabulary Learning for University Students</b> Ismail Xodabande, Yasaman Iravi, Behzad Mansouri and Hoda Matinparsa
50	<b>The Impact of Mobile-Assisted Language Learning on English as a Foreign Language Learners' Vocabulary Learning Attitudes and Self-Regulatory Capacity</b> Xiao Lei, Jalil Fathi, Shabnam Noorbakhsh and Masoud Rahimi
64	<b>Using Mobile Devices for Vocabulary Learning Outside the Classroom: Improving the English as Foreign Language Learners' Knowledge of High-Frequency Words</b> Azadeh Rahmani, Vahid Asadi and Ismail Xodabande
71	<b>A Corpus Study of Lexical Bundles Used Differently in Dissertations Abstracts Produced by Chinese and American PhD Students of Linguistics</b> Kai Bao and Meihua Liu
84	<b>Integrating Automatic Speech Recognition Technology Into Vocabulary Learning in a Flipped English Class for Chinese College Students</b> Michael Yi-Chao Jiang, Morris Siu-Yung Jong, Na Wu, Bin Shen, Ching-Sing Chai, Wilfred Wing-Fat Lau and Biyun Huang
100	<b>Developing morphological knowledge with online corpora in an ESL vocabulary classroom</b> Rui Zhang
113	<b>Incidental vocabulary acquisition from listening to English teacher education lectures: A case study from Macau higher education</b> Barry Lee Reynolds, Xiaowen (Serina) Xie and Quy Huynh Phu Pham
131	<b>Learning English vocabulary from word cards: A research synthesis</b> Yuanying Lei and Barry Lee Reynolds





## OPEN ACCESS

EDITED AND REVIEWED BY  
Xiaolin Zhou,  
Peking University, China

\*CORRESPONDENCE  
Ehsan Rassaei  
ehsanrassaei@yahoo.com

SPECIALTY SECTION  
This article was submitted to  
Language Sciences,  
a section of the journal  
Frontiers in Psychology

RECEIVED 27 October 2022  
ACCEPTED 31 October 2022  
PUBLISHED 20 January 2023

CITATION  
Rassaei E (2023) Editorial:  
Incorporating technology for teaching  
vocabulary to second language  
learners. *Front. Psychol.* 13:1081901.  
doi: 10.3389/fpsyg.2022.1081901

COPYRIGHT  
© 2023 Rassaei. This is an  
open-access article distributed under  
the terms of the [Creative Commons  
Attribution License \(CC BY\)](#). The use,  
distribution or reproduction in other  
forums is permitted, provided the  
original author(s) and the copyright  
owner(s) are credited and that the  
original publication in this journal is  
cited, in accordance with accepted  
academic practice. No use, distribution  
or reproduction is permitted which  
does not comply with these terms.

# Editorial: Incorporating technology for teaching vocabulary to second language learners

Ehsan Rassaei\*

Majan University College, Muscat, Oman

## KEYWORDS

vocabulary, technology, call, L2, smartphones, language learning

## Editorial on the Research Topic

### Incorporating technology for teaching vocabulary to second language learners

Since learning vocabulary plays an essential role in second language (L2) acquisition, a lot of attention has been paid to developing techniques as well as investigating factors that lead to and promote L2 vocabulary learning (e.g., Nassaji and Tian, 2010; Huang and Lin, 2014; Sun, 2017; Rassaei, 2018; Nguyen and Boers, 2019). The bulk of the extant research on L2 vocabulary has been devoted to investigating techniques that are beneficial to L2 vocabulary learning. These studies highlighted the importance of several factors which should be considered in designing L2 vocabulary learning tasks. One such factor is the level of L2 learners' engagement with the learning task. According to the involvement load hypothesis (Hulstijn and Laufer, 2001), the level of learners' cognitive involvement with the vocabulary learning task, which is determined by three factors namely search, choice, and evaluation, is essential for vocabulary learning. An implication of this hypothesis is that vocabulary learning tasks that are more cognitively engaging are more effective for L2 vocabulary learning. Further research also indicated the benefits of productive vocabulary learning activities that engage learners in recycling new words productively while performing learning tasks (e.g., Laufer, 2006; Keating, 2008; Min, 2008; Rassaei, 2017). These studies in general suggest that activities that result in more elaborative processing lead to more fruitful results.

Meanwhile, as L2 acquisition researchers have become increasingly more interested in applying different forms of technology in language learning, interest in incorporating technology to teach vocabulary has been also growing in recent years. An increasing number of researchers have become interested in investigating the effects of different forms of technology in vocabulary learning (e.g., Lin and Yu, 2017; Rassaei, 2020; Teng and Zhang, 2021; Lo, 2022; Muñoz et al., 2022). Several features of Web-2 technology such as sharing information and interaction in real time which promote collaborative learning made various forms of technology appealing to SLA researchers. Moreover, the availability of portable devices such as smartphones which provided further affordances

for language learning has sparked an increasing number of studies with regard to vocabulary learning (e.g., Lin and Yu, 2017; Xu and Peng, 2017). Various technology tools such as computer applications, virtual spaces and other digital resources have made collaborative language learning easier and more successful among L2 learners (Gánem-Gutiérrez, 2018).

In line with the current interest in applying various forms of technology for vocabulary learning and instruction, the present article Research Topic aimed to collect and present a number of studies that investigated how technology can be used to teach vocabulary to L2 learners. As the first study in this Research Topic, Soleimani et al. investigated how smartphones can provide assistance for vocabulary learning. In their study, a group of EFL learners read a number of short texts that included some unfamiliar vocabulary items. Another group of participants read the same texts and listened to the audio file of the same texts through their smartphones. The results revealed the value of smartphones in enhancing L2 vocabulary knowledge. The findings also indicated the benefits of smartphones in increasing learners' vocabulary learning self-efficacy. The effectiveness of smartphones for L2 vocabulary learning has also been investigated by other studies in the present article Research Topic. Polakova and Kimova examined the effects of a mobile application that provided several affordances for L2 learners including translation, text to speech and pronunciation functions. The results revealed that those students who used the application improved in vocabulary learning and experienced higher motivation and satisfaction for vocabulary learning.

In a longitudinal study, Lei et al. examined the effects of mobile-assisted vocabulary learning on EFL learners' attitudes toward vocabulary learning and their self-regulation. The results revealed the positive effects of mobile-assisted vocabulary learning on learners' vocabulary learning attitudes and their self-regulation scores. These three studies in the present Research Topic provided evidence for the value and effectiveness of smartphones in enhancing L2 learners' vocabulary knowledge as well as promoting L2 learners' attitudes toward vocabulary learning.

The affordances provided by smartphones for L2 vocabulary learning have also been examined by two other papers in the present Research Topic. Xodabande et al. investigated how digital flashcards, installed on the participants' smartphones, can be used for teaching academic words to university students. The results indicated that the participants who used digital flashcards outperformed the group that used traditional paper flashcards as well as the control group. Rahmani et al. also investigated how a digital flashcard application installed on smartphones could enhance EFL learners' vocabulary knowledge. The results were promising as those students who used the application outperformed the control group that was exposed to regular language learning activities in the absence of the digital flashcard. The results of the studies published in the present

Research Topic and reviewed above, point to the benefits of smartphones and the relevant applications installed on them for teaching vocabulary to L2 learners.

Turning to other studies published in this Research Topic, Jiang et al. investigated the effects of automatic speech recognition (ASR) technology on vocabulary learning in a tertiary flipped setting. The results indicated that the group who incorporated ASR technology into their classroom activities outperformed the control group that did not use ASR during the treatment session in terms of vocabulary learning.

Two studies in the present Research Topic adopted a research synthesis approach to investigate the use of technology for teaching vocabulary to L2 learners. Wei and Fan reviewed the studies that investigated the effects of different forms of on-screen texts such as subtitles and captions on L2 vocabulary learning. The authors reported L2 captions to be more effective than (L1) subtitles. The authors also reported that some individual learner factors can also moderate the effectiveness of on-screen texts for vocabulary learning. Lei and Reynolds also synthesized research on learning L2 English vocabulary from word cards based on 32 primary studies. The authors reported a larger effect size for studies that investigated paper word cards than digital word cards and also for the ready-made word cards compared with self-constructed word cards. One reason for why learners benefited more from paper word cards compared to digital word cards can be attributed to learners' familiarity with paper word cards and their lack of experience to work with digital word cards.

In another study, Reynolds et al. investigated the impact of listening to audio-recorded lectures on incidental vocabulary learning. The learner participant of the study listened to the audio-recorded lectures provided via Google Drive for about 14 weeks. The findings indicated the benefits of listening to the lectures for incidental vocabulary learning. The student also indicated positive attitudes toward listening to the lectures for vocabulary learning. Finally, two studies in this Research Topic investigated the affordances provided by corpora for L2 vocabulary learning. Bao and Liu examined how corpora can be used to compare the differences in the use of lexical bundles between Chinese and American students writing their Ph.D. dissertations' abstracts. The findings indicated major variations in the use of lexical bundles in both groups. The results provided evidence for the importance of using corpora for identifying lexical bundles which pose difficulty for language learners. Finally, Zhang explored the use of online corpora to enhance learners' morphological knowledge and L2 vocabulary learning. The findings indicated the conducting effects of using corpora in language classrooms to enhance vocabulary learning.

In sum, the articles published in this Research Topic referred the affordances that technology can provide for L2 vocabulary learning. In particular, these studies indicated the benefits of mobile-assisted language learning in L2

vocabulary learning. Further technology affordances for L2 vocabulary learning as indicated in the present Research Topic include on-screen texts, digital flashcards, audio-recorded lectures, digital corpora and speech recognition technology. These studies provided evidence for the values and the effectiveness of different forms of technology. Moreover, the studies indicate a variety of technology tools that can be utilized in various contexts depending on factors such as learning and teaching styles as well as their availability. In addition to the evidence provided by studies in the present Research Topic regarding the usefulness of the technology tools in enhancing L2 collaborative vocabulary learning, one implication which can be inferred from these studies is the diversity and variety of the technology tools that can be adapted by language teachers to suit specific purposes in specific contexts.

## Author contributions

The author confirms being the sole contributor of this work and has approved it for publication.

## References

- Gánem-Gutiérrez, G. A. (2018). "Collaborative activity in the digital world," in *The Routledge Handbook of Sociocultural Theory and Second Language Development*, eds J. P. Lantolf, M. E. Poehner, with M. Swain (London: Routledge), 391–408. doi: 10.4324/9781315624747-25
- Huang, L. L., and Lin, C. C. (2014). Three approaches to glossing and their effects on vocabulary learning. *System* 44, 127–136. doi: 10.1016/j.system.2014.03.006
- Hulstijn, J. H., and Laufer, B. (2001). Some empirical evidence for the involvement load hypothesis in vocabulary acquisition. *Lang. Learn.* 51, 539–558. doi: 10.1111/0023-8333.00164
- Keating, G. D. (2008). Task effectiveness and word learning in a second language: the involvement load hypothesis on trial. *Lang. Teach. Res.* 12, 365–386. doi: 10.1177/1362168808089922
- Laufer, B. (2006). Comparing focus on form and focus on forms in second-language vocabulary learning. *Can. Mod. Lang. Rev.* 63, 149–166. doi: 10.3138/cmlr.63.1.149
- Lin, C. C., and Yu, Y. C. (2017). Effects of presentation modes on mobile-assisted vocabulary learning and cognitive load. *Interact. Learn. Environ.* 25, 528–542. doi: 10.1080/10494820.2016.1155160
- Lo, S. (2022). Learning vocabulary through dual-subtitled viewing: the impact of different ILH-based interventions. *Comput. Assist. Lang. Learn.* 2022, 1–28. doi: 10.1080/09588221.2022.2126497
- Min, H. T. (2008). EFL vocabulary acquisition and retention: reading plus vocabulary enhancement activities and narrow reading. *Lang. Learn.* 58, 73–115. doi: 10.1111/j.1467-9922.2007.00435.x
- Muñoz, C., Pattenmore, A., and Avello, D. (2022). Exploring repeated captioning viewing as a way to promote vocabulary learning: time lag between repetitions and learner factors. *Comput. Assist. Lang. Learn.* 2022, 1–27. doi: 10.1080/09588221.2022.2113898
- Nassaji, H., and Tian, J. (2010). Collaborative and individual output tasks and their effects on learning English phrasal verbs. *Lang. Teach. Res.* 14, 397–419. doi: 10.1177/1362168810375364
- Nguyen, C. D., and Boers, F. (2019). The effect of content retelling on vocabulary uptake from a TED talk. *TESOL Quart.* 53, 5–29. doi: 10.1002/tesq.441
- Rassaei, E. (2017). Effects of three forms of reading-based output activity on L2 vocabulary learning. *Lang. Teach. Res.* 21, 76–95. doi: 10.1177/1362168815606160
- Rassaei, E. (2018). Computer-mediated textual and audio glosses, perceptual style and L2 vocabulary learning. *Lang. Teach. Res.* 22, 657–675. doi: 10.1177/1362168817690183
- Rassaei, E. (2020). Effects of mobile-mediated dynamic and nondynamic glosses on L2 vocabulary learning: a sociocultural perspective. *Mod. Lang. J.* 104, 284–303. doi: 10.1111/modl.12629
- Sun, C. H. (2017). The value of picture-book reading-based collaborative output activities for vocabulary retention. *Lang. Teach. Res.* 21, 96–117. doi: 10.1177/1362168816655364
- Teng, M. F., and Zhang, D. (2021). The associations between working memory and the effects of multimedia input on L2 vocabulary learning. *Int. Rev. Appl. Linguist. Lang. Teach.* 2021, 130. doi: 10.1515/iral-2021-0130
- Xu, Q., and Peng, H. (2017). Investigating mobile-assisted oral feedback in teaching Chinese as a second language. *Comput. Assist. Lang. Learn.* 30, 173–182. doi: 10.1080/09588221.2017.1297836

## Acknowledgments

I would like to thank the authors who submitted their papers to this Research Topic. Also, I appreciate Prof. Andy Gao for his assistance in preparing this Research Topic.

## Conflict of interest

The author declares that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

## Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.



# Exploring the Effect of Assisted Repeated Reading on Incidental Vocabulary Learning and Vocabulary Learning Self-Efficacy in an EFL Context

Habib Soleimani, Farnoosh Mohammaddokht and Jalil Fathi\*

Department of English and Linguistics, Faculty of Language and Literature, University of Kurdistan, Sanandaj, Iran

## OPEN ACCESS

### Edited by:

Ehsan Rassaei,  
Majan University College, Oman

### Reviewed by:

Mostafa Azari Noughabi,  
University of Gonabad, Iran  
Nasim Ghanbari,  
Persian Gulf University, Iran

### \*Correspondence:

Jalil Fathi  
Jfathi13@yahoo.com  
orcid.org/0000-0003-1146-1024

### Specialty section:

This article was submitted to  
Language Sciences,  
a section of the journal  
Frontiers in Psychology

**Received:** 10 January 2022

**Accepted:** 31 January 2022

**Published:** 18 February 2022

### Citation:

Soleimani H, Mohammaddokht F and  
Fathi J (2022) Exploring the Effect of  
Assisted Repeated Reading on  
Incidental Vocabulary Learning and  
Vocabulary Learning Self-Efficacy in  
an EFL Context.  
Front. Psychol. 13:851812.  
doi: 10.3389/fpsyg.2022.851812

The purpose of the current study was to investigate the effect of two types of repeated reading (i.e., assisted and unassisted) on incidental vocabulary learning of Iranian English as a Foreign Language (EFL) learners. In so doing, a sample of 45 intermediate EFL students from two intact classes of a language institute were selected as the participants. The two classes were randomly assigned to an unassisted group ( $N=21$ ) who were required to just *read* and an assisted group ( $N=24$ ) who were asked to *read* and *listen* to 24 short texts several times. The assisted group employed their smartphones to listen to the audio files of the short stories. The data were gathered *via* a researcher-made vocabulary test and vocabulary learning self-efficacy scale. The results of ANCOVA revealed that although both types of repeated reading contributed to enhancing vocabulary learning of the participants, assisted repeated reading led to significantly greater EFL vocabulary gains. Additionally, the findings revealed that both assisted and unassisted repeated reading improved vocabulary learning self-efficacy of the participants and there was not a significant difference between the two types of interventions. The findings of the present study have implications for EFL researchers and practitioners.

**Keywords:** repeated reading, assisted reading, incidental vocabulary learning, vocabulary learning self-efficacy, smartphone

## INTRODUCTION

Vocabulary learning is long regarded as an essential aspect of second and foreign language (L2/FL) learning (Nation, 2013). Through vocabulary learning, L2 learners can achieve mastery over the second language (Nation, 2008). Since practitioners and scholars have come to realize the integral role of vocabulary learning in communication and language learning, more research attention has been given to L2 vocabulary instruction (Rassaei, 2017; Yousefi and Biria, 2018; Liu et al., 2020). L2 literature has observed a growing interest in exploring effective explicit and implicit strategies such as using dictionaries, inferring word meaning from context, and extensive reading for learning collocations and vocabulary in particular (see Hunt and Beglar, 2005). Research has also shown that a considerable amount of vocabulary is learned receptively

through listening or reading (Nagy et al., 1985). In the meantime, due to growing popularity of new technologies and applications, numerous researchers and practitioners have investigated and employed mobile assisted language learning (MALL) as a viable technique for L2 vocabulary teaching and learning (Burston, 2013; Rassaei, 2018, 2020; Lin and Lin, 2019).

Given that reading has been considered as a critical source of vocabulary growth (Zahar et al., 2001), the role of reading for L2 vocabulary learning should receive further attention. What is commonly agreed upon is that some part of one's L2 vocabulary is acquired incidentally through reading (Ramos and Dario, 2015). One of the most influential procedures within L2 education investigations is extensive reading (Nakanishi, 2015) which has been emphasized to develop vocabulary learning (Wang, 2013). Extensive reading is a type of reading that provides learners with exposure to large quantities of materials (Grabe and Stoller, 2002) for comprehension often without performing any tasks after reading. Extensive reading has been continuously reported as the most commonly recommended model improving learners' language proficiency such as reading comprehension (Nakanishi, 2015), grammar knowledge (Ellis, 2005) and incidental vocabulary learning (Horst, 2005; Suk, 2017). Nonetheless, in order for the learning to occur, it is worth stressing that the reading materials need to be selected according to the learners' language proficiency and reading abilities (Suk, 2017). When learners do multiple readings, they have rapid access to known words and word patterns in various contexts through encountering them repeatedly. Over the time, learners' vocabulary size tends to develop and they can also achieve a deeper understanding of the new words. Words which are learned in this procedure can be incorporated into learners' writing and speech (Nation, 2008).

Investigations into incidental vocabulary learning have revealed that the number of times a new and unknown word is repeated in a text influences how likely individuals will learn the word successfully (e.g., Peters and Webb, 2018). In this regard, delving deeply into vocabulary development, a number of empirical studies have been carried out to explore the effectiveness of repeated reading method in the past decades, hoping to find out how incidental vocabulary learning can occur through this approach (see Webb and Chang, 2012). In the repeated reading approach, the students read and reread a passage several times (two to four times) aloud (Samuels, 1979) or silently (Anderson, 1993) in a predetermined level of pace until reading the text fluently (Meyer and Felton, 1999). In this approach, learners read specified texts from graded readers (e.g., books and passages that have simplified grammatical structures and reduced vocabulary range) repeatedly in order to develop word recognition as well as improving reading fluency and comprehension (Dlugosz, 2000). There are two kinds of repeated reading: assisted and unassisted repeated reading procedures. Unassisted repeated reading is where students read short passages autonomously without any audiotape to follow until they reach a fluent reading state. With the same procedure, in assisted form of repeated reading students read along while listening to the audiotape or a live model (Webb and Chang, 2012). Samuels who first coined the term repeated reading pinpointed

that unassisted repeated reading method increases poor readers' comprehension and oral fluency (Samuels, 1976, 1979). In addition, Chomsky (1978), found that the repeated reading approach caused slow readers to become more motivated, confident, and willing to read new materials independently.

Repeated reading has been regarded as one of the most effective approaches for acquiring vocabulary because it can expose L2 students to massive amounts of meaningful input. Therefore, these multiple and consistent exposures and repetition contribute to the incidental acquisition of novel English vocabulary (Suk, 2017). With regard to the EFL context, some researchers have highlighted the effective role of assisted repeated reading in vocabulary gains (Webb and Chang, 2012; Liu and Todd, 2016; Serrano and Huang, 2018) and fluency development (Taguchi et al., 2004). According to Serrano and Huang (2018), students are able to learn vocabulary even when the assisted repeated reading focuses on comprehension. Furthermore, Taguchi et al. (2004) argued that assisted repeated reading has the potential to develop readers' fluency as well as helping them become independent readers.

Concerning self-efficacy, Bandura (1977, 1978) believed that individuals with high assurance in their skills and capabilities can succeed in performing a difficult task and see it as a challenge to be mastered not as a threat to be avoided. One's self-efficacy determines not only their persistence, endeavor, and strategizing, but their subsequent job performance (Heslin and Klehe, 2006). Some educators so far have adapted self-efficacy and categorized it further, for instance, into professional self-efficacy, multitasking self-efficacy, and computer self-efficacy, and (Islam et al., 2018). Among other dimensions of self-efficacy, in this study we decomposed self-efficacy into vocabulary self-efficacy. As far as EFL context is concerned, skill-specific self-efficacy has recently received some attention in empirical studies (e.g., Fathi et al., 2019, 2020; Fathi and Soleimani, 2020; Rahimi and Fathi, 2021).

Although there is a plethora of studies investigating vocabulary learning through reading as well as the impressive progress which has been made by L2 researchers (see Horst, 2005; Webb, 2007; Webb and Chang, 2012), some gaps still exist. First, research lacks a comprehensive and conclusive theory of in what ways incidental vocabulary is learned through repeated reading in an EFL context. Although Webb and Chang (2012) and Serrano and Huang (2018) underscored the effectiveness of assisted and unassisted repeated reading in vocabulary learning among Taiwanese EFL learners, further studies should be carried out in EFL contexts. Second, repeated reading (i.e., assisted and unassisted) research provides relatively little evidence about the incidental vocabulary-expanding impacts of reading repeatedly, simply because L2 literature mostly tends to focus on more general aspects of language development. Furthermore, given the widespread recognition of MALL (Burston, 2013), the use of smartphones for assisted repeated reading has remained under-researched. Finally, to the best of our knowledge, no study to date has explored the effect of repeated reading, including its both forms on vocabulary learning self-efficacy among EFL learners. The rationale behind investigating vocabulary learning self-efficacy was the fact that



self-efficacy is argued to influence vocabulary learning and the use of appropriate strategies for vocabulary learning (Mizumoto, 2012, 2013; Hong et al., 2014). In addition, although L2 vocabulary has received significant research attention, self-efficacy in vocabulary learning has remained under-explored (Mizumoto, 2013; Nation, 2013). Given the insufficient available data, further investigation is required to determine whether and exactly how repeated reading leads to vocabulary learning. In an attempt to bridge the gap in the current literature, this study investigated the scope and depth of incidental vocabulary learning and vocabulary learning self-efficacy through assisted and unassisted repeated reading in the EFL setting of Iran. Therefore, the following research questions guided this research:

1. Does assisted repeated reading have any significant effect on incidental vocabulary learning of EFL learners?
2. Does assisted repeated reading have any significant effect on vocabulary learning self-efficacy of EFL learners?

## REVIEW OF LITERATURE

It is commonly thought that vocabulary learning is pivotal in L2/FL learning (e.g., Nation, 2013). Language learners might face an immense vocabulary challenge since they require to learn different dimensions of each word, including its aural and textual forms, collocations and associations (Nation, 2013), in order to successfully understand and use L2 communicatively. Researchers have also stressed the significance of word knowledge in reading comprehension (see Read, 2000). It has been argued that a learner who knows more than 98% of the words used in a passage can fully understand it autonomously (e.g., Schmitt et al., 2011). Therefore, it can be argued that a good vocabulary knowledge including both its breadth and depth is a significant element of reading competencies.

Cronbach's (1942) conceptualization categorized vocabulary knowledge into two major divisions: first, knowledge of word meaning (generalization, precision and breadth of meaning) and second, degrees of accessibility to this knowledge (application and availability). Later, Richards (1976) introduced other factors such as syntax, register, frequency, association, derivation, polysemy, and semantic features involved in knowing a word. Qian (1999) also proposed two key aspects of vocabulary knowledge: Vocabulary breadth knowledge that is related to the size of vocabulary and depth of vocabulary knowledge that is referred to how deep and well one knows a word.

The extant literature on L2 learning provides much empirical credit to the significance of vocabulary in the language acquisition process (Read, 1988). To date, some topics related to vocabulary in EFL/L2 have been investigated. In particular, the vocabulary knowledge of EFL learners has attracted increasing attention (Rassaei, 2017; Yousefi and Biria, 2018). Some studies have examined EFL learners' vocabulary learning through explicit instruction (Mizumoto and Takeuchi, 2009), computer-assisted language learning (Shokrpour et al., 2019; Namaziandost et al., 2021), extensive reading (Liu and Zhang, 2018; Song, 2020), and repeated reading (Liu and Todd, 2016). Considering the

fact that vocabulary learning process is complex and explicit vocabulary teaching in EFL classrooms could only cover a small proportion of new words that students learn, it is important to find other ways for teaching vocabulary. What can be found in the current literature is some mentions of instructional techniques for enhancing L2 vocabulary learning reported by different instructors, such as the learning of word lists (Carter, 1987), learning words in a discourse context (Laufer, 2003), inferencing (Alahmadi and Foltz, 2020), exposure to word glosses (Webb, 2007), songs (Pavia et al., 2019), games or stories (Chou, 2014). The detailed accounts of the implementation of these techniques in real EFL classrooms and their effectiveness in vocabulary teaching have been the concern of many researchers. Teaching vocabulary seems laborious for EFL teachers since teaching English is very likely to encounter many obstacles and challenges. As Gorsuch et al. (2015) mentioned, one challenge for FL programs is to provide students with sufficient input and experience to use the language. There are few opportunities for EFL learners to use English in the real world since they have little or even no exposure to the language beyond the classroom walls (Read, 1988).

Scholars and practitioners have become growingly cognizant of the salient effect of reading, more particularly extensive reading on vocabulary knowledge (Horst, 2005). Extensive reading refers to a kind of L2 reading approach in which (a) learners are provided with a considerable number of reading materials with the purpose of reading for pleasure; (b) they read their selection at a fast rate to obtain a general understanding of the text; (c) they are more related with comprehending the whole passage rather than individual sentences or words (Day et al., 1998). Intensive reading, in contrast, often is concerned with the careful and precise reading of more difficult, shorter FL passages with the purpose of detailed and thorough comprehension under the guidance of the teacher (Carrell and Carson, 1997).

Nation and Wang (1999) asserted that graded readers can be seen as one of the key sources of vocabulary learning for L2 learners if used appropriately. As extensive reading approach provides students with the chance to face different vocabularies in their context of use, it can be very pleasant and motivating which in turn can facilitate learner autonomy (Thornbury, 2002). A bulk of L1 and L2/FL reading research has mentioned the potential merits of extensive reading in expanding language learners' vocabulary learning (Pigada and Schmitt, 2006; Liu and Wu, 2011). One example of such investigations is Pigada and Schmitt's (2006) study which investigated the effect of extensive reading on vocabulary. The study tested 133 words in a month with a learner of French as a FL. The results showed that nearly two-thirds of the target words were learned. This is to say, extensive reading can be an effective approach in promoting incidental vocabulary learning. In a subsequent study, Liu and Zhang (2018) did a meta-analysis study in this field of research. Their first goal was to discover the role of extensive reading in vocabulary learning. The second aim was to find appropriate teaching methods and teaching length in order to have an effective extensive reading program. The findings of their study indicated that extensive reading led to

significantly greater vocabulary knowledge (English vocabulary). They also suggested that extensive reading can develop learners' vocabulary in one semester (less than 3 months). In addition, different comprehension questions, vocabulary exercises, and graded readers are appropriate and influential teaching methods and reading materials for improving EFL learners' vocabulary promoting.

In a similar vein, the results from Kweon and Kim's (2008) study demonstrated that the participants achieved vocabulary gains after the EFL extensive reading intervention. It was also revealed that the participants found nouns easier to retain compared to verbs and adjectives. Furthermore, they learned frequent words more easily than less frequent words. Moreover, Horst's (2005) pilot study indicated that the participants learned mostly half of the unfamiliar words they faced in the extensive reading materials they selected. In this line of research, Wang (2013) made an attempt to examine the effect of extensive reading on the word knowledge of EFL Taiwanese learners. This was a 15-week extensive reading procedure in which the participants were required to read 30 English texts in this period. The findings underscored the significant impact of extensive reading on English vocabulary growth. It was claimed that the EFL extensive reading program plays a beneficial role in incidental vocabulary learning among EFL learners with lower competence. However, implementing extensive reading without the supervision of the teacher outside the classroom may lead students to face difficulties (Martina et al., 2020).

In the last few years, L2 researchers and scholars have assembled an intriguing list of research probing repeated reading (several readings of a passage) as an approach for enhancing reading competencies (e.g., Gorsuch and Taguchi, 2010). Although there is still much to be learned about this procedure, evidence shows that repeated reading is a viable instructional model for both disabled and developmental readers (Therrien, 2004).

Rereading the same text using either the assisted or unassisted repeated reading procedure significantly enhances reading rate and accuracy (Chomsky, 1978; Samuels, 1979). In repeated reading students read a meaningful passage repeatedly until oral production is flowing and fluid and results in increased comprehension and fluency (Taguchi, 1997). Basically, repeated reading falls into two categories: assisted repeated reading (*read-along*), in which a student reads the passage while audiotaped or live model of the text is used (Chomsky, 1978). And unassisted repeated reading (*independent practice*), where the student reads the passage autonomously while no model or prototype is used (Samuels, 1979). Samuels (1979) devised the repeated reading approach as means of supporting incompetent readers achieve automatic word recognition. In this procedure, students were required to reread a meaningful passage aloud until reaching a criterion degree of fluency.

Repeated reading has so far drawn increasing attention from researchers as a potentially effective method for reading and comprehension among L1 (Kuhn and Stahl, 2003) and L2 readers (Webb and Chang, 2012; Gorsuch et al., 2015). According to Blum and Koskinen (1991) the positive effect of repeated reading on reading fluency has been revealed to have other considerable benefits for learners. They maintained that repeated

reading not only improves reading comprehension and fluency, but also helps learners become more motivated to read and more confident in their reading.

Repeated reading is also a method which scaffolds students with reading disabilities to build fluency (Therrien and Hughes, 2008; Lee and Yoon, 2017). For example, students with reading disabilities in the study carried out by Lee and Yoon (2017) showed high reading fluency after the repeated reading intervention. They also stressed the effectiveness of the combination of repeated reading and a listening passage preview for students with reading disabilities, more specifically for those at the elementary level. Furthermore, Therrien (2004) found that repeated reading is an effective procedure for both students with learning disabilities and nondisabled students by which their reading fluency and comprehension increases.

The related literature evinces that audio-assisted repeated reading is a promising procedure for improving L2 readers' fluency (Taguchi et al., 2004). Repeated reading of meaningful passage, including listening-while reading (assisted form) has been found to produce improvements in reading fluency, rate and word recognition accuracy (Rasinski, 1990). The findings from Taguchi et al.'s (2004) study yielded that repeated reading helped students to learn and retain vocabulary and grammar. Students were also able to monitor their reading comprehension through multiple readings. It was also revealed that neither successive re-readings nor audio-models were tedious and distracting for a considerable number of participants.

As far as L2 vocabulary learning is considered, many scholars have corroborated the effectiveness of repeated reading for vocabulary growth (Serrano and Huang, 2018). Drawing on insights acquired from investigations into repeated reading as well as on reading aloud, some researchers believe that repeated encounters with new forms in various contexts make repeated reading instrumental for incidental L2 vocabulary learning (Brown et al., 2008). Incidental vocabulary learning research has verified the assumption that much exposure to L2 reading texts can contribute to vocabulary growth (Webb and Chang, 2012).

In this line of research, several studies have explored the incidental vocabulary learning through repeated reading (Brown et al., 2008; Webb and Chang, 2012; Liu and Todd, 2016; Serrano and Huang, 2018). For instance, the main focus of the study conducted by Webb and Chang (2012) was on the influence of repeated reading on incidental vocabulary learning. The findings revealed that both kinds of repeated reading (assisted and unassisted) contributed to improvement in incidental vocabulary learning. More precisely, the Taiwanese learner of English as a FL reported significant word gains through assisted repeated reading.

Similarly, Serrano and Huang (2018) explored the impact of assisted form of repeated reading on vocabulary learning. Taiwanese EFL students as participants of this study were divided into two groups of intensive and spaced. The participants in one group were provided with an assisted repeated reading program (reading while listening) once every day (intensive) while the learners in the other were supposed to read the same passage once every week (spaced). The results showed

that vocabulary gains were achieved after the assisted repeated reading treatment. It is worth mentioning that the participants in the intensive group exhibited significantly greater vocabulary achievement compared to the spaced group. This suggests that concentrated practice contributes to greater vocabulary learning.

More recently, Serrano and Huang (2018) set out a partial replication study of Serrano and Huang's (2018) previous research into incidental vocabulary learning. This study used the same context, methodology, design, and analyses as the original study; the level of the target vocabulary learning was the only different factor. The results showed a significant intentional vocabulary gains after the repeated reading interval. The only common finding between two studies refers to immediate vocabulary gains. However, Liu and Todd (2016) argued that repeated reading seems more effective with target words that have etymological roots with the learners' L1.

In recent years, the role of non-cognitive factors of language learning, such as self-efficacy motivation, and emotions has come under the spotlight. In a general sense, self-efficacy refers to individuals' perceptions towards their ability in accomplishing a specific task successfully (Bandura, 2011; Fathi et al., 2021). The important non-cognitive trait influences achievement, skills, knowledge, perceived value, and outcome expectations (Schunk, 2003). As Linnenbrink and Pintrich (2003) argued, efficacious learners are more likely to persist, seek help, and work hard, so they can successfully complete a task. Extending this concept into the domain of reading, self-efficacy is the reader's belief about their capability to read effectively (Guthrie and Wigfield, 1999). According to Wigfield et al. (2004), self-efficacious readers have not only better performance but also tend to persist through difficult reading tasks.

The focus of the current study is the effect of repeated reading on learners' self-efficacy in learning words, vocabulary self-efficacy more specifically. Less is known about the association between self-efficacy and repeated reading, because previous studies have mainly focused on the effect of repeated reading on reading comprehension. To the best knowledge of researchers, no study to date has investigated the impact of repeated reading on vocabulary self-efficacy in the EFL context.

In summary, the review of the existing literature indicated that incidental vocabulary learning through repeated reading (i.e., assisted and unassisted) has not been seriously taken into consideration by the practitioners in the domain of EFL, in the context of Iran in particular. Unfortunately, no studies so far have systematically and explicitly investigated incidental English vocabulary learning in Iran. The data regarding vocabulary and vocabulary mastery in EFL, however, is not sufficient to give a vivid picture of the role of repeated reading since vocabulary growth seems to be neglected. Therefore, little information is available with regard to the rationale behind EFL instructors' adoption of repeated reading approach for incidental vocabulary learning. In order to shed more light on this procedure, the current study examined the effect of assisted and unassisted repeated reading on incidental vocabulary learning and sought to determine whether and to what extent lexical knowledge of the participants was developed.

## MATERIALS AND METHODS

### Participants

A total number of 45 Iranian EFL students took part in this quasi-experimental study. These participants were pre-intermediate students of two intact groups in a private language institute in Tehran, Iran. The two classes were randomly assigned to an assisted repeated group ( $N=21$ ) and an unassisted repeated group ( $N=24$ ). All the participants were female students ranging in age from 13 to 16, with the mean age of 14.56. Also, they had the experience of at least 4 to 6 years of learning English as a foreign language either in public schools or in the private language teaching institutes. The global English proficiency of the participants was measured by a version of Preliminary English Test (PET). The mean scores of PET were compared by running an independent samples *t*-test whose results indicated that there was no statistically significant difference between the groups.

### Reading Materials

The reading materials were 26 short stories selected from <http://eslyes.com/eslread/>. Much care was exercised to choose stories of similar level of difficulties. **Table 1** presents selected short stories and their linguistics features. As seen in the table, Flesch–Kincaid Grade Level of all stories ranged from 3 to 3.6. Their ease score varied from 85.9 to 93.1. Most texts lacked passive sentences and the texts had from minimum of 8.6 to maximum of 12.1 average number of words per sentence. The total number of words varied from 192 to 320. In addition to the selected 26 short stories, the students were required to study *Top Notch 2*, a famous commercially published series, as the major textbook prescribed by the institute.

### Materials and Instruments

#### Preliminary English Test

Before beginning the treatment, the homogeneity of the students concerning global English proficiency was checked. To this end, a version of Preliminary English Test (PET) by Cambridge English for Speakers of Other Languages (ESOL, 2009) was given to the students of both groups. This PET version comprised three components including Reading (5 sections), Listening (4 sections), and Speaking (4 sections). The reliability coefficients of the reading and listening components were 0.83 and 0.79, respectively. Also, the inter-rater reliability coefficient for the speaking component was reported to be 0.81.

#### Vocabulary Test

Vocabulary learning of the participants was assessed by a 70-item multiple choice test designed by the researchers. The items of this test were randomly chosen from the short stories which were provided to both groups. The stems of the items were selected from the statements of the stories and the distractors were also the vocabularies included in content of the short stories. Two parallel forms of the test were designed for the pre-test and post-test by altering the order of items and distractors. The face and content validity of the test were approved by three domain experts. In addition,



**TABLE 1** | Selected short stories and their linguistics features.

	Short story title	Flesch–Kincaid grade level	Flesch reading ease score	%Passive sentences	Average number of words per sentence	Total number of words
1	Man flies 200 miles in chair	3.0	93.1	0%	11.3	226
2	Cleaning a dirty plate—oops!	3.4	88.7	0%	10.5	316
3	Toilet tank almost overfills	3.4	87.7	0%	9.7	282
4	A play, or a movie? neither!	3.2	91.6	4%	11.7	247
5	You're not my dad	3.5	88.2	0%	10.8	305
6	A haircut every 2 weeks	3.3	92.5	0%	12.1	219
7	Yardman mows and blows	3.3	91.4	0%	12.1	209
8	Finds bargains at thrift shop	3.3	89.4	4%	10.3	227
9	Pete's too sharp knife	3.2	87.7	4%	8.9	223
10	A good hot dog sandwich	3.5	87.5	0%	10.0	211
11	Loose button gets sewn	3.5	88.8	4%	10.7	225
12	Paper pile grows and goes	3.1	90.3	0%	10.1	192
13	A daytime robbery in LA	3.5	86.0	0%	9.3	214
14	English is so hard	3.2	90.4	0%	11.1	201
15	Let us buy some paint	3.5	86.9	0%	9.7	283
16	A big cash wedding gift	3.3	88.7	3%	10.0	291
17	New to America from Asia	3.2	87.0	0%	8.6	199
18	Am i having a heart attack?	3.4	87.3	3%	10.0	320
19	Scrooge brings christmas gift	3.3	88.0	0%	11.1	211
20	His haircut leaves her cold	3.2	88.7	0%	9.5	209
21	Driving lesson scares them both	3.4	87.0	0%	9.2	304
22	God: open door, feed flies	3.3	90.8	0%	11.5	230
23	He: fraud; She: little white lie	3.2	89.0	0%	10.2	316
24	Man hoards library books	3.4	89.1	3%	10.6	278
25	TSA revises jewelry regs	3.4	85.9	13%	9.5	219
26	Cancer? white spot on tongue	3.6	86.8	6%	10.0	312

a pilot study was performed to gain preliminary information regarding the reliability and appropriateness of the items. The internal consistency of the test, as estimated by KR-21, was 0.79.

### Vocabulary Learning Self-Efficacy Scale

Vocabulary learning self-efficacy of the participants was measured by the 4 items adapted from Mizumoto (2013). Each item was assessed on a 6-point scale varying from 1 (not at all true of me) to 6 (very true of me). A sample item of the scale is “I am good at memorizing vocabulary.” The reliability estimate of this scale, as calculated by Cronbach's Alpha formula, was reported to be 0.78 in this study.

### Procedure

This experiment began in summer semester of 2019. During the first session, the researcher administered the PET as well as the pre-tests of the study which included the vocabulary test and VLSS. The purpose of this course was to improve the general language proficiency of the participants. However, in addition to *Top Notch 2*, the students of both groups were provided with short stories as the supplementary materials. The two groups were taught by the same instructor at a private language institute. Following the procedure used by Webb and Chang (2012), the researchers began the study intervention. The intervention which lasted for 13 sessions was carried out to investigate the effects of repeated reading of 26 short stories over a 13 session period. Both groups

were required to read two short stories in each session. More specifically, the unassisted group students were required just to read the stories, whereas the students of assisted group were required to read and listen to the stories. The assisted group students were also provided with audio files of the stories in addition to the texts. The assisted group employed their smartphones to listen to the audio files of the short stories. However, the students of the unassisted group were just provided with the printed version of stories without audio files and they were not required to use their smartphones as the course requirement. The students of both groups were required to either read (i.e., unassisted) or read and listen (i.e., assisted) to each story at least three times. The purpose of both reading conditions was comprehension of the texts without any focus on any particular grammatical structure or vocabulary explanation. The students could make use of dictionary, discuss the texts with each other, or ask questions. At the end of the treatment, the post-tests (vocabulary test and VLSS) were given to the participants of both groups.

### Data Analysis

Both descriptive and inferential statistics were employed for the data analysis. Concerning the inferential data analysis, the collected data were analyzed by performing one-way between-groups analyses of covariance (ANCOVA). The purpose of running ANCOVA was to compare the impacts of the two types of reading instructions (i.e., assisted vs. unassisted) used in the two groups on vocabulary learning

**TABLE 2** | Results of the PET for each group.

Groups	M (SD)	T	Sig.
Assisted group	25.16 (8.28)	−0.621	0.436
Unassisted group	27.08 (9.01)		

**TABLE 3** | Descriptive statistics for pre- and post-tests scores.

Groups	Scales	Pre-test		Post-test	
		M	SD	M	SD
Assisted group	Vocabulary	65.33	11.85	81.04	12.84
	Self-efficacy	2.64	0.54	3.07	0.59
Unassisted group	Vocabulary	64.12	10.32	72.20	11.49
	Self-efficacy	2.77	0.58	3.27	0.47

and vocabulary learning self-efficacy as the two dependent variables of the study.

## RESULTS

To ensure the homogeneity of the participants in terms of global English skills, an independent-samples *t*-test was performed to compare the PET scores for the assisted and unassisted groups. As observed in **Table 2**, the outcomes showed that no significant difference was observed in the PET scores for the assisted group ( $M=25.16$ ,  $SD=8.28$ ) and the unassisted group [ $M=27.08$ ,  $SD=9.01$ ;  $t(43)=-0.621$ ,  $p>0.05$ ], confirming the fact that both groups were homogeneous before beginning the intervention.

Afterwards, in order to explore the impacts of the two types of repeated reading on the participants' vocabulary learning and vocabulary learning self-efficacy, a number of one-way between-groups analysis of covariance (ANCOVA) were conducted to compare the effects of the two types of L2 reading instructions (i.e., assisted vs. unassisted repeated reading) used in the two groups on the two dependent variables.

Concerning the impact of using the repeated reading on EFL learners' vocabulary learning, as indicated in **Table 3**, the vocabulary learning mean score of the assisted group was 65.33 ( $SD=11.85$ ) on the pre-test and it was increased to 81.04 ( $SD=12.84$ ) on the post-test. By the same token, the mean score of vocabulary learning on the pre-test for the unassisted group was raised from 64.12 ( $SD=10.32$ ) to 72.20 ( $SD=11.49$ ) on the post-test. However, after adjusting for the pre-test scores of vocabulary learning, it was revealed that there was a statistically significant difference between the two groups on post-test scores of vocabulary learning, [ $F(1, 42)=7.97$ ,  $p=0.007$ , partial  $\eta^2=0.16$ ; see **Table 4**]. This finding demonstrated that the participants of the assisted group improved their vocabulary learning significantly more than the participants of the unassisted group, demonstrating that the assisted repeated reading instruction was significantly more effective in enhancing the L2 vocabulary learning of the participants.

With regard to vocabulary learning self-efficacy, the descriptive statistics data (see **Table 3**) indicate that the mean score of the vocabulary learning self-efficacy for the unassisted group was 2.77 ( $SD=0.58$ ) in the pre-test and it was raised to 3.27 ( $SD=0.47$ ) on the post-test. Similarly, the vocabulary learning self-efficacy mean score for the assisted group was 2.64 ( $SD=0.54$ ) on the pre-test and this value was raised to 3.07 ( $SD=0.59$ ) on the post-test. After adjusting for the pre-test scores of vocabulary learning self-efficacy, the results of ANCOVA (see **Table 5**) showed that there was not any statistically significant difference between the two groups on post-test scores of vocabulary learning self-efficacy, [ $F(1, 42)=1.04$ ,  $p=0.312$ , partial  $\eta^2=0.02$ ]. This finding revealed that both assisted and unassisted repeated reading instructions equally enhanced the vocabulary learning self-efficacy of the participants and there was not any significant difference between them.

## DISCUSSION

The aim of this research was to explore the effects of assisted and unassisted repeated reading on incidental vocabulary learning among Iranian EFL learners. Furthermore, the role of both kinds of the repeated reading in influencing vocabulary learning self-efficacy was examined. The results from this study offer some key findings: First, repeated reading proved to boost vocabulary learning. It was found that both groups (assisted and unassisted) manifested significant improvements in terms of vocabulary learning after the intervention. However, assisted repeated reading helped learners to gain substantially further EFL words. This finding accords with that of Webb and Chang (2012) which revealed that assisted and unassisted repeated reading enhanced vocabulary knowledge of Taiwanese EFL learners with assisted repeated reading contributed to further vocabulary gains. This finding also supports those of Brown et al. (2008) who indicated that reading while listening (assisted repeated reading) is more useful than either listening or reading solely for incidental vocabulary learning. This finding also re-echoes Serrano and Huang's (2018) claim that assisted repeated reading promotes vocabulary learning in EFL contexts. Since the assisted group used smartphone as the technology device, this finding is also partially in line with previous studies (e.g., Burston, 2013; Rassaei, 2018, 2020; Lin and Lin, 2019) which have emphasized the influence of technology for vocabulary learning.

The findings revealed that vocabulary knowledge was enhanced significantly from multiple exposures to an unknown word – learners seemingly *picked-up* the words. Three interpretations seem valid in this regard. One possible explanation may be that repeated reading coupled with audiotapes provided students with practice in automaticity of vocabulary recognition. Each new encounter with a word during repeated reading might have been conducive to enrich learners' knowledge of that word through repetition. It can be argued that since the students had more opportunity to re-read the passages, they had greater potential to learn the words incidentally. Put differently, greater opportunity for repetition helped them to consolidate the

**TABLE 4 |** The results of ANCOVA on vocabulary size.

Source	Type III sum of squares	df	Mean square	F	Sig.	Partial $\eta^2$
Corrected model	3460.320	2	1730.160	19.348	0.000	0.480
Intercept	1193.185	1	1193.185	13.343	0.001	0.241
Pre.vocabulary	2585.230	1	2585.230	28.911	0.000	0.408
Group	713.199	1	713.199	7.976	0.007	0.160
Error	3755.680	42	89.421			
Total	269421.000	45				
Corrected total	7216.000	44				

**TABLE 5 |** The results of ANCOVA on vocabulary learning self-efficacy.

Source	Type III sum of squares	df	Mean square	F	Sig.	Partial $\eta^2$
Corrected model	7.519	2	3.759	30.719	0.000	0.594
Intercept	2.804	1	2.804	22.911	0.000	0.353
Pre.Self-efficacy	7.077	1	7.077	57.834	0.000	0.579
Group	0.128	1	0.128	1.049	0.312	0.024
Error	5.140	42	0.122			
Total	467.208	45				
Corrected total	12.658	44				

knowledge of novel and partly known words. This interpretation is supported by some researchers, suggesting that greater frequency of word meeting contributes to word learning (Laufer and Rozovski-Roitblat, 2011).

Second, as the words were met repeatedly, learners could guess their meanings from the context. That is to say, when words are repeated the exposures of the words after the first encounter may provide an opportunity for guessing meanings from the context clues, facilitating retrieval of the meanings of the words obtained from previous meetings. Learners also had the opportunity to learn deliberately through using a dictionary for looking up the meanings of words as there was no time limitation in this procedure.

Third, greater learning may have taken place in the current study due to the fact that the assisted repeated reading group was given the opportunity to read the same passages multiple times while listening to the words being repeated. This suggests that the visual (reading the written passage) and the phonological (listening to the audiotapes) may double the chance of the words being incidentally learned. This means that the passages were more likely to be less effective in enhancing incidental vocabulary learning without re-reading and listening simultaneously. Aural support while reading has been argued to have a positive impact on L2 vocabulary learning (e.g., Brown et al., 2008; Webb and Chang, 2012). It can be claimed that the prosodic characteristics of assisted repeated reading might have aided EFL learners in splitting the linguistic data in chunks more meaningfully, which might have contributed to global understanding of the text. The better global text comprehension, in turn, has enhanced students' competence in guessing the unfamiliar words.

The second finding of this study was that both kinds of repeated reading contributed to self-efficacy in vocabulary

learning among the EFL learners and no significant difference was found between the two types of treatments. Learners' self-motivated attention to words facilitates their vocabulary learning, which in turn results in greater self-efficacy. Further exposure of both groups to vocabularies included in the short stories might have helped them to enhance their confidence in vocabulary learning. Following Bandura's (1977) notion of self-efficacy, it can be argued that repeated reading might have boosted learners' beliefs in their own competencies in directing and executing the learners' activities and strategies in learning L2 vocabularies. Additionally, since self-efficacy is argued to be correlated with self-regulation (Zimmerman and Martinez-Pons, 1990; Su et al., 2019), it is likely that the two kinds of repeated reading have boosted students' self-regulation in vocabulary learning, which in turn has contributed to improving students' vocabulary learning self-efficacy.

Overall, the results obtained from this study provide two findings that inform practice. First, results revealed that both assisted and unassisted repeated reading can be effective approaches for enhancing the amount of L2 incidental vocabulary learning while there can be much to gain from assisted repeated reading procedure which leads to significantly greater vocabulary learning. Repeated reading approach can be used effectively in order to improve learners' ability to learn vocabulary incidentally. Additionally, both forms of repeated reading increased vocabulary self-efficacy among the EFL learners. The current study findings can expand the existing literature regarding vocabulary learning and repeated reading as it lent empirical support to the usefulness of repeated reading instruction in influencing incidental vocabulary learning.

Assisted repeated reading (in which students listen over and over to an audiotape while reading) helps students

comprehend the passage and encourage deeper insights. Additional rereading may help the learners remember more meaningful structures, increase their reading accuracy, and *pick up* the new words. As a result, learners may show greater understanding and use the words mentioned in the text unconsciously while talking about the passage. If a passage is read a number of times, learners benefit substantially by remembering more words.

## IMPLICATIONS

In view of the results of this study, some pedagogical recommendations are suggested to EFL instructors. EFL practitioners are recommended to incorporate smartphones and other applications into their classrooms for vocabulary teaching. As class hours might not suffice for teachers to teach vocabularies effectively, technology devices can help students to learn vocabularies outside the class in their convenient time. Besides, technology devices and applications can not only add fun and excitement to learning but also help EFL students become self-regulated learners (Guirguis and Antigua, 2017). Additionally, practitioners play an influential role in ensuring the usefulness of the repeated reading approach. Therefore, it is needed for future teachers to be trained in executing repeated reading instruction and its techniques such as assessment of learners' vocabulary level, the selection of appropriate reading materials, and design of reading assignments given to learners. Before implementing assisted or unassisted repeated reading, learners should have a thorough understanding of repeated reading theory, such as benefits, principles, and the ways to make repeated reading successful. Teachers should explain the difference between assisted and unassisted repeated reading and various strategies for developing vocabulary learning before learners start reading.

It is recommended that EFL teachers keep passages short (about 50 to 300 words) because long passages may tire students. They can take various passages from different types of reading materials such as newspaper articles, magazines, novels, and short stories commensurate with learners' proficiency levels. Reading materials selected for repeated reading program should address learners' interests and learning needs so as to motivate and energize them to read the passages. It is important not to choose a passage that contains too many unknown words because it might be demanding for learners to comprehend and guess unfamiliar words from the context. If the aim of repeated reading procedure is to enable students to learn words through a particular passage, learners should be given a cue to focus on words and the text should be repeated more than three times.

## LIMITATIONS AND SUGGESTIONS FOR FURTHER RESEARCH

The findings of the current research, nonetheless, have some limitations. To provide a more in-depth and more conclusive

evaluation of repeated reading, it is suggested that future studies use qualitative methods in addition to tests or questionnaires in order to present a more detailed and holistic image of the role of repeated reading in improving incidental vocabulary learning. For example, semi-structured interviews about learners' perceptions towards assisted and unassisted repeated reading would add novel insights into the current literature. Moreover, this study used the data from EFL learners from a private institute in the context of Iran. This context might be radically different compared to other contexts (either private or public schools) in affecting students' vocabulary learning. Future studies are needed to examine other samples of EFL learners from both the private and public sectors. Furthermore, the existing finding from this sample of Iranian EFL learners may not be transferable to other EFL contexts and cultures. Further research is required to achieve details in association with repeated reading in other populations. In spite of the fact that the outcomes show that repeated reading is useful for learners, the majority of studies have not taken learners' reading levels into account. Therefore, the success of assisted and unassisted repeated reading for learners with various reading levels cannot be identified. Also, the role of other individual differences in affecting the usefulness of repeated reading should be investigated. Finally, this research did not examine vocabulary retention to find how long such vocabulary learning might last. As a result, it is undefined whether the learned vocabularies were yet retained in the participants' mental lexicon after the procedure had been finished. Future research thus needs to employ a delayed assessment to investigate the impact of repeated reading on the longer-term retention of acquired vocabularies. Finally, to enrich our understanding regarding repeated reading, exploring the impacts of assisted and unassisted repeated reading on learners' reading motivation, reading anxiety, and reading attitude would be interesting topics for future studies.

## DATA AVAILABILITY STATEMENT

The raw data supporting the conclusions of this article will be made available by the authors, without undue reservation.

## ETHICS STATEMENT

The studies involving human participants were reviewed and approved by University of Kurdistan. The patients/participants provided their written informed consent to participate in this study.

## AUTHOR CONTRIBUTIONS

HS, FM, and JF were equally involved in designing the research, topic development, data collection, data analysis, writing drafts, and final editing. All authors contributed to the article and approved the submitted version.



## REFERENCES

- Alahmadi, A., and Foltz, A. (2020). Effects of language skills and strategy use on vocabulary learning through lexical translation and inferencing. *J. Psycholinguist. Res.* 49, 975–991. doi: 10.1007/s10936-020-09720-9
- Anderson, N. (1993). "Repeated reading," in *New Ways in Teaching Reading*. ed. R. R. Day (Alexandria, VA: Teachers of English to Speakers of Other Languages), 190–191.
- Bandura, A. (1977). Self-efficacy: toward a unifying theory of behavioral change. *Psychol. Rev.* 84, 191–215. doi: 10.1037/0033-295X.84.2.191
- Bandura, A. (1978). Reflections on self-efficacy. *Adv. Behav. Res. Ther.* 1, 237–269. doi: 10.1016/0146-6402(78)90012-7
- Bandura, A. (2011). On the functional properties of perceived self-efficacy revisited. *J. Manag.* 38, 9–44. doi: 10.1177/0149206311410606
- Blum, I. H., and Koskinen, P. S. (1991). Repeated reading: a strategy for enhancing fluency and fostering expertise. *Theory Pract.* 30, 195–200.
- Brown, R., Waring, R., and Donkaewbua, S. (2008). Incidental vocabulary acquisition from reading, reading-while-listening, and listening to stories. *Nat. Foreign Lang. Resour. Center* 20:816.
- Burston, J. (2013). Mobile-assisted language learning: a selected annotated bibliography of implementation studies 1994–2012. *Lang. Learn. Technol.* 17, 157–225.
- Carrell, P. L., and Carson, J. G. (1997). Extensive and intensive reading in an EAP setting. *Engl. Specif. Purp.* 16, 47–60. doi: 10.1016/S0889-4906(96)00031-2
- Carter, R. (1987). *Vocabulary: Applied Linguistic Perspectives*. London: Allen and Unwin.
- Chomsky, C. (1978). "When you still can't read in third grade. After decoding, what?" in *What Research Has to Say about Reading Instruction*. ed. S. Samuels (Newark, Del: International Research Association), 13–30.
- Chou, M. H. (2014). Assessing English vocabulary and enhancing young English as a foreign language (EFL) learners' motivation through games, songs, and stories. *Education* 42, 284–297. doi: 10.1080/03004279.2012.680899
- Cronbach, L. J. (1942). An analysis of techniques for diagnostic vocabulary testing. *J. Educ. Res.* 36, 206–217. doi: 10.1080/00220671.1942.10881160
- Day, R. R., Bamford, J., Renandya, W. A., Jacobs, G. M., and Yu, V. W. S. (1998). Extensive reading in the second language classroom. *REL C J.* 29, 187–191. doi: 10.1177/003368829802900211
- Dlugosz, D. W. (2000). Rethinking the role of reading in teaching a foreign language to young learners. *Eng. Lang. Teach. J.* 54, 284–290. doi: 10.1093/elt/54.3.284
- Ellis, R. (2005). Principles of instructed language learning. *System* 33, 209–224. doi: 10.1016/j.system.2004.12.006
- ESOL (2009). *Cambridge BEC Preliminary: Official Examination Papers from University of Cambridge ESOL Examinations*. England: Cambridge University Press.
- Fathi, J., Ahmadnejad, M., and Yousofi, N. (2019). Effects of blog-mediated writing instruction on L2 writing motivation, self-efficacy, and self-regulation: a mixed methods study. *J. Res. Appl. Ling.* 10, 159–181. doi: 10.22055/RALS.2019.14722
- Fathi, J., Derakhshan, A., and Torabi, S. (2020). The effect of listening strategy instruction on second language listening anxiety and self-efficacy of Iranian EFL learners. *SAGE Open* 10:878. doi: 10.1177/2158244020933878
- Fathi, J., Greenier, V., and Derakhshan, A. (2021). Self-efficacy, reflection, and burnout among Iranian EFL teachers: the mediating role of emotion regulation. *Iranian J. Lang. Teach. Res.* 9, 13–37. doi: 10.30466/IJLTR.2021.121043
- Fathi, J., and Soleimani, H. (2020). The effect of reading strategy instruction on reading self-efficacy and reading attitudes: a case of young female Iranian EFL learners. *Appl. Res. Eng. Lang.* 9, 382–408. doi: 10.22108/ARE.2019.116944.1461
- Gorsuch, G., and Taguchi, E. (2010). Developing reading fluency and comprehension using repeated reading: evidence from longitudinal student reports. *Lang. Teach. Res.* 14, 27–59. doi: 10.1177/1362168809346494
- Gorsuch, G., Taguchi, E., and Umehara, H. (2015). Repeated reading for Japanese language learners: effects on reading speed, comprehension, and comprehension strategies. *Read. Matrix* 15, 18–44.
- Grabe, W., and Stoller, F. L. (2002). *Teaching and Researching Reading*. Harlow, UK: Longman.
- Guirguis, R., and Antigua, K. C. (2017). DLLs and the development of self-regulation in early childhood. *Cogent Educ.* 4:1355628. doi: 10.1080/2331186X.2017.1355628
- Guthrie, J. T., and Wigfield, A. (1999). How motivation fits into a science of reading. *Sci. Stud. Read.* 3, 199–205. doi: 10.1207/s1532799xssr0303\_1
- Heslin, P. A., and Klehe, U. C. (2006) in *Encyclopedia of Industrial/Organizational Psychology*. Vol. 2. ed. S. G. Rogelberg (California: SAGE), 705–708.
- Hong, J. C., Hwang, M. Y., Tai, K. H., and Chen, Y. L. (2014). Using calibration to enhance students' self-confidence in English vocabulary learning relevant to their judgment of over-confidence and predicted by smartphone self-efficacy and English learning anxiety. *Comput. Educ.* 72, 313–322. doi: 10.1016/j.compedu.2013.11.011
- Horst, M. (2005). Learning L2 vocabulary through extensive reading: a measurement study. *Can. Mod. Lang. Rev.* 61, 355–382. doi: 10.3138/cmlr.61.3.355
- Hunt, A., and Beglar, D. (2005). A framework for developing EFL reading vocabulary. *Read. Foreign Lang.* 17, 23–59.
- Islam, A. K. M. N., Mavengere, N. B., Ahlfors, U.-R., Ruohonen, M. J., Serenko, A., and Palvia, P. (2018). A stress-strain-outcome model of job satisfaction: the moderating role of professional self-efficacy [Conference Paper]. *24th Americas Conference on Information Systems*, New Orleans, USA.
- Kuhn, M. R., and Stahl, S. A. (2003). Fluency: a review of developmental and remedial practices. *J. Educ. Psychol.* 95, 3–21. doi: 10.1037/0022-0663.95.1.3
- Kweon, S. O., and Kim, H. R. (2008). Beyond raw frequency: incidental vocabulary acquisition in extensive reading. *Read. Foreign Lang.* 20, 191–215.
- Laufer, B. (2003). Vocabulary acquisition in a second language: do learners really acquire most vocabulary by reading? *Canadian Mod. Lang. Rev.* 59, 567–587. doi: 10.3138/cmlr.59.4.567
- Laufer, B., and Rozovski-Roitblat, B. (2011). Incidental vocabulary acquisition: The effects of task type, word occurrence and their combination. *Lang. Teach. Res.* 15, 391–411. doi: 10.1177/1362168811412019
- Lee, J., and Yoon, S. Y. (2017). The effects of repeated reading on reading fluency for students with reading disabilities: a meta-analysis. *J. Learn. Disabil.* 50, 213–224. doi: 10.1177/0022219415605194
- Lin, J. J., and Lin, H. (2019). Mobile-assisted ESL/EFL vocabulary learning: a systematic review and meta-analysis. *Comput. Assist. Lang. Learn.* 32, 878–919. doi: 10.1080/09588221.2018.1541359
- Linnenbrink, E. A., and Pintrich, P. R. (2003). The role of self-efficacy belief in student engagement and learning in the classroom. *Read. Writing Q.* 19, 119–137. doi: 10.1080/10573560308223
- Liu, L., Akhter, S., and Qureshi, A. H. (2020). Towards the description of techniques in teaching L2 vocabulary. *Revista Argentina de Clínica Psicológica* 29:268.
- Liu, Y. T., and Todd, A. G. (2016). Implementation of assisted repeated reading techniques for the incidental acquisition of novel foreign vocabulary. *Lang. Teach. Res.* 20, 53–74. doi: 10.1177/1362168814559802
- Liu, H., and Wu, H. (2011). Impact of extensive reading on young EFL learners' vocabulary knowledge, reading proficiency, and motivation for learning English. *NPUST Human. Soc. Sci. Res.* 5, 1–21.
- Liu, J., and Zhang, J. (2018). The effects of extensive Reading on English vocabulary learning: a meta-analysis. *Engl. Lang. Teach.* 11, 1–15. doi: 10.5539/elt.v11n6p1
- Martina, F., Syafryadin, S., and Utama, J. A. (2020). The practice of extensive reading among EFL learners in tertiary level. *Yavana Bhasha: J. Eng. Lang. Educ.* 3, 56–72. doi: 10.25078/yb.v3i2.1712
- Meyer, M. S., and Felton, R. H. (1999). Repeated reading to enhance fluency: old approaches and new directions. *Ann. Dyslexia* 49, 283–306. doi: 10.1007/s11881-999-0027-8
- Mizumoto, A. (2012). Exploring the effects of self-efficacy on vocabulary learning strategies. *Stud. Self-Access Learn. J.* 3, 423–437. doi: 10.37237/030407
- Mizumoto, A. (2013). Effects of self-regulated vocabulary learning process on self-efficacy. *Innov. Lang. Learn. Teach.* 7, 253–265. doi: 10.1080/17501229.2013.836206
- Mizumoto, A., and Takeuchi, O. (2009). Examining the effectiveness of explicit instruction of vocabulary learning strategies with Japanese EFL university students. *Lang. Teach. Res.* 13, 425–449. doi: 10.1177/1362168809341511
- Nagy, W. E., Herman, P. A., and Anderson, R. C. (1985). Learning words from context. *Read. Res. Q.* 20, 233–253. doi: 10.2307/747758
- Nakanishi, T. (2015). A meta-analysis of extensive reading research. *TESOL Q.* 49, 6–37. doi: 10.1002/tesq.157
- Namazandost, E., Alekasir, S., Dehkordi, E. S., and Tilwani, S. A. (2021). An account of EFL learners' vocabulary learning in a mobile-assisted language environment: the case of Rosetta stone application. *Computer-Assist. Lang. Learn. Elect. J.* 22, 80–110.

- Nation, I. S. P. (2008). *Teaching Vocabulary: Strategies and Techniques*. Boston: Heinle.
- Nation, I. S. P. (2013). *Learning Vocabulary in Another Language*. 2nd Edn. (England: Cambridge University Press).
- Nation, I. S. P., and Wang, K. (1999). Graded readers and vocabulary. *Read. Foreign Lang.* 12, 355–380.
- Pavia, N., Webb, S., and Faez, F. (2019). Incidental vocabulary learning through listening to songs. *Stud. Second. Lang. Acquis.* 41, 745–768. doi: 10.1017/S0272263119000020
- Peters, E., and Webb, S. (2018). Incidental vocabulary acquisition through viewing L2 television and factors that affect learning. *Stud. Second. Lang. Acquis.* 40, 551–577. doi: 10.1017/S0272263117000407
- Pigada, M., and Schmitt, N. (2006). Vocabulary acquisition from extensive reading: A case study. *Read. Foreign Lang.* 18, 1–28.
- Qian, D. D. (1999). Assessing the roles of depth and breadth of vocabulary knowledge in reading comprehension. *Canadian Mod. Lang. Rev.* 56, 282–308. doi: 10.3138/cmlr.56.2.282
- Rahimi, M., and Fathi, J. (2021). Exploring the impact of wiki-mediated collaborative writing on EFL students' writing performance, writing self-regulation, and writing self-efficacy: a mixed methods study. *Comput. Assist. Lang. Learn.*, 1–48. doi: 10.1080/09588221.2021.1888753
- Ramos, R., and Dario, F. (2015). Incidental vocabulary learning in second language acquisition: A literature review. *Profile Issues Teac. Prof. Dev.* 17, 157–166. doi: 10.15446/profile.v17n1.43957
- Rasinski, T. V. (1990). Effects of repeated reading and listening-while-reading on reading fluency. *J. Educ. Res.* 83, 147–151. doi: 10.1080/00220671.1990.10885946
- Rassaei, E. (2017). Effects of three forms of reading-based output activity on L2 vocabulary learning. *Lang. Teach. Res.* 21, 76–95. doi: 10.1177/1362168815606160
- Rassaei, E. (2018). Computer-mediated textual and audio glosses, perceptual style and L2 vocabulary learning. *Lang. Teach. Res.* 22, 657–675. doi: 10.1177/1362168817690183
- Rassaei, E. (2020). Effects of mobile-mediated dynamic and nondynamic glosses on L2 vocabulary learning: a sociocultural perspective. *Mod. Lang. J.* 104, 284–303. doi: 10.1111/modl.12629
- Read, J. (1988). Measuring the vocabulary knowledge of second language learners. *RELJ* 19, 12–25. doi: 10.1177/003368828801900202
- Read, J. (2000). *Assessing Vocabulary*. Cambridge: Cambridge University Press.
- Richards, J. C. (1976). The role of vocabulary teaching. *TESOL Q.* 10, 77–89. doi: 10.2307/3585941
- Samuels, S. J. (1976). Automatic decoding and reading comprehension. *Lang. Arts* 53, 323–325.
- Samuels, S. J. (1979). The method of repeated reading. *Read. Teach.* 32, 403–408.
- Schmitt, N., Jiang, X., and Grabe, W. (2011). The percentage of words known in a text and reading comprehension. *Mod. Lang. J.* 95, 26–43. doi: 10.1111/j.1540-4781.2011.01146.x
- Schunk, D. H. (2003). Self-efficacy for reading and writing: influence of modeling, goal setting, and self-evaluation. *Read. Writing Q.* 19, 159–172. doi: 10.1080/10573560308219
- Serrano, R., and Huang, H. Y. (2018). Learning vocabulary through assisted repeated reading: how much time should there be between repetitions of the same text? *TESOL Q.* 52, 971–994. doi: 10.1002/tesq.445
- Shokrpour, N., Mirshekari, Z., Moslehi, S., and Popescu, M. (2019). Learning vocabulary electronically: does computer assisted language learning (CALL) instruction have any impacts on Iranian EFL learners? *Cogent Educ.* 6:1702827. doi: 10.1080/2331186X.2019.1702827
- Song, M. (2020). The impacts of extensive Reading on EFL primary school students' vocabulary acquisition and Reading comprehension. *J. Exten. Read.* 5, 60–69.
- Su, Y., Li, Y., Liang, J.-C., and Tsai, C.-C. (2019). Moving literature circles into wiki-based environment: the role of online self-regulation in EFL learners' attitude toward collaborative learning. *Comput. Assist. Lang. Learn.* 32, 556–586. doi: 10.1080/09588221.2018.1527363
- Suk, N. (2017). The effects of extensive reading on reading comprehension, reading rate, and vocabulary acquisition. *Read. Res. Q.* 52, 73–89. doi: 10.1002/rrq.152
- Taguchi, E. (1997). The effects of repeated readings on the development of lower identification skills of FL readers. *Read. Foreign Lang.* 11, 97–119.
- Taguchi, E., Takayasu-Maass, M., and Gorsuch, G. J. (2004). Developing reading fluency in EFL: how assisted repeated reading and extensive reading affect fluency development. *Read. Foreign Lang.* 16, 70–96.
- Therrien, W. J. (2004). Fluency and comprehension gains as a result of repeated reading: a meta-analysis. *Remedial Spec. Educ.* 25, 252–261. doi: 10.1177/07419325040250040801
- Therrien, W. J., and Hughes, C. (2008). Comparison of repeated reading and question generation on students' reading fluency and comprehension. *Learn. Disabil. Contemp. J.* 6, 1–16.
- Thornbury, S. (2002). *How to Teach Vocabulary*. Essex: Pearson Education Limited.
- Wang, Y. (2013). Incidental vocabulary learning through extensive reading: A case of lower-level EFL Taiwanese learners. *J. Asia TEFL* 10, 59–80.
- Webb, S. (2007). Learning word pairs and glossed sentences: the effects of a single context on vocabulary knowledge. *Lang. Teach. Res.* 11, 63–81. doi: 10.1177/1362168806072463
- Webb, S., and Chang, A. C. (2012). Vocabulary learning through assisted and unassisted repeated reading. *Canadian Mod. Lang. Rev.* 68, 267–290. doi: 10.3138/cmlr.1204.1
- Wigfield, A., Guthrie, J. T., Tonks, S., and Perencevich, K. C. (2004). Children's motivation for reading: domain specificity and instructional influences. *J. Educ. Res.* 97, 299–310. doi: 10.3200/JOER.97.6.299-310
- Yousefi, M. H., and Biria, R. (2018). The effectiveness of L2 vocabulary instruction: a meta-analysis. *Asian-Pacific J. Second Foreign Lang. Educ.* 3, 1–19. doi: 10.1186/s40862-018-0062-2
- Zahar, R., Cobb, T., and Spada, N. (2001). Acquiring vocabulary through reading: effects of frequency and contextual richness. *Canadian Mod. Lang. Rev.* 57, 541–572. doi: 10.3138/cmlr.57.4.541
- Zimmerman, B. J., and Martinez-Pons, M. (1990). Student differences in self-regulated learning: relating grade, sex, and giftedness to self-efficacy and strategy use. *J. Educ. Psychol.* 82, 51–59. doi: 10.1037/0022-0663.82.1.51

**Conflict of Interest:** The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

**Publisher's Note:** All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Copyright © 2022 Soleimani, Mohammaddokht and Fathi. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.



# Vocabulary Mobile Learning Application in Blended English Language Learning

Petra Polakova\* and Blanka Klimova

Department of Applied Linguistics, Faculty of Informatics and Management, University of Hradec Králové, Hradec Králové, Czechia

## OPEN ACCESS

### Edited by:

Xuesong Gao,  
University of New South Wales,  
Australia

### Reviewed by:

Lucas Kohnke,  
The Education University of  
Hong Kong, Hong Kong SAR, China  
Jalil Fathi,  
University of Kurdistan, Iran

### \*Correspondence:

Petra Polakova  
petra.polakova.3@uhk.cz

### Specialty section:

This article was submitted to  
Language Sciences,  
a section of the journal  
Frontiers in Psychology

Received: 03 February 2022

Accepted: 21 March 2022

Published: 03 May 2022

### Citation:

Polakova P and Klimova B (2022)  
Vocabulary Mobile Learning  
Application in Blended English  
Language Learning.  
Front. Psychol. 13:869055.  
doi: 10.3389/fpsyg.2022.869055

Mobile devices and applications, which have become an integral part of our lives, are gradually used for different purposes, including learning languages in EFL classrooms. Since vocabulary plays an important role in the process of foreign language learning, the aim of this study was to explore the use of the vocabulary mobile learning application and its usefulness in blended English learning. Quantitative and qualitative approach to research was used, since the integration of both approaches created the possibility to solve complex research problem. The case study was based on the use of the developed mobile application called *Angličtina Today* the content of which corresponded to the language needs of the target group of students. The quantitative approach used a method of quasi-experiment aiming to achieve the pre-tests and post-test results of the students from the experimental and control groups. The results showed that the students facing blended learning, including mobile application in the process of language learning, achieved better results than the students exposed to the traditional, face-to-face education. In addition, the results revealed students' overall satisfaction with the application. The main reasons for their satisfaction were improved vocabulary knowledge, ease of use, and enhanced motivation. Based on these findings from the current study, it can be argued that the vocabulary mobile learning application proved to be useful in the process of blended English language learning.

**Keywords:** mobile applications, apps, English language, vocabulary learning, blended learning

## INTRODUCTION

In recent years, blended learning (BL) has become popular in educational environment. As it combines traditional and online learning modes, the promise of BL rests on the strengths of both learning approaches. Having this in consideration, the most important aim of the BL design is to find an effective combination of different learning methods which can motivate students to participate inside and outside of the classroom environment (Neumeier, 2005; Senffner and Kepler, 2015). Senffner and Kepler (2015) pointed out that BL is flexible, scalable, and meaningful way of learning since its online component allows students to learn anytime and anywhere without being limited to groups or partners. Regarding this, the mobile technology appears to be an appropriate approach to education since smartphones and other portable Wi-Fi gadgets are related to both traditional and innovative ways of learning and are well

aligned with strategic educational goals (Kukulska-Hulme, 2009). Klímová (2017) argued that characteristics such as portability, individuality, unobtrusiveness, availability, adaptability, persistence, usefulness, and usability make mobile devices an ideal language-learning tool. Moreover, the ownership of smartphones among students has grown fast which led to a great popularity of the mobile technology in education, including English as Foreign Language (EFL) learning. Mobile devices have been adapted in the traditional classroom environment and are increasingly used by English language learners (Nasab and Taki, 2016).

Since vocabulary acquisition is considered to be an essential part of language learning, vocabulary mobile learning applications (apps) have become a popular form of mobile-assisted language learning (MALL) (Klímová, 2019). Research has shown that a mobile application designed and based on students' needs and continuously facilitated by a teacher is effective in the enhancement of students' performance and contributes to positive learning outcomes (Mahdi, 2017; Rezaei et al., 2013; Klímová, 2019). Despite an increasing number of mobile learning applications offering a powerful learning environment, it is certain that the traditional methods of learning foreign languages are still useful. However, they need to be updated and combined with new trends in education (Mikulecký, 2019). Before the onset of coronavirus, in Slovakia, teaching and learning had been mainly based on the traditional face-to-face learning approach. The technology-based learning tools were implemented only in the case when teachers were willing to do so. After the spread of coronavirus, the traditional learning was disrupted and sudden lockdown forced teachers and students to adapt their teaching and learning. New circumstances made educational institution recognize the importance of technology-based learning tools since they appear to be effective tools for ensuring students' education. This is especially true for mobile devices. Considering the importance of mobile technology, the present study investigated a mobile learning application called *Angličtina Today* [English Today] and its usefulness in the process of vocabulary acquisition in the EFL setting of Slovakia. The following research questions guided this research:

1. To what extent is the use of a mobile application useful for second language vocabulary acquisition?
2. What are the perceptions of EFL students of the vocabulary mobile learning application?
3. Has the use of the mobile application helped students to extend their vocabulary knowledge?

## LITERATURE REVIEW

It is commonly thought that vocabulary lies in the center of language learning since it is impossible to convey information without having knowledge of vocabulary (Wilkins, 1972; Nation, 2001; Zhang, 2011; Alquahtani, 2015; Gürkan, 2018). In order to increase vocabulary size, both incidental and instructional vocabulary acquisition are required. Incidental

learning has been portrayed as implicit since knowledge is acquired independently of conscious attempts to learn. In other words, incidental learning is taken as it is learning without intent to learn (Nation, 2001; Eraut, 2004). According to Schmitt (2010), only ten encounters with vocabulary can lead to its sizeable gain since the repetition affects incidental vocabulary learning. According to his findings, learners who encounter an unknown word more times in informative contexts are able to demonstrate significantly larger gains in vocabulary knowledge than learners who have fewer encounters with a new word. This is in line with Soleimani et al. (2022) whose study revealed that repeated exposure to vocabulary leads to significant EFL vocabulary gains. Despite of this, there is an unknown boundary between word which has been recognized and the word which has been fully understood and learnt in the process of incidental vocabulary learning. Therefore, there is a need of some systematic approach, to ensure vocabulary acquisition which can be reached by intentional learning (Björgvinsson, 2012). It is in accordance with Zhang and Teng (2012) whose study pronounced a significant role of intentional learning. Students need to get input in forms of explanation and clear instructions and they need to practice the language they have learnt which could possibly lead to a successful output (Birkner, 2016). The role of input, information processing, and output has been gradually acknowledged in second language learning. Zhang (2009) believed that these are essential elements in foreign language acquisition, including vocabulary acquisition. Different phases of vocabulary learning can take place in different environments such as traditional school learning environment or technology-based learning environment. Combination of educational technology with traditional methods can lead to effective vocabulary acquisition and thus to achievement of balance in the learning process (Mahdi, 2017).

Poláková (2019) recommended the combination of online interaction with traditional classroom methods. She assumed that blended learning can make learning process more interesting and motivational. Moreover, the previous studies reported positive attitudes of the students toward the use of the mobile application in learning vocabulary which is in accordance with the findings of the study conducted by Lu (2008) who found out that the positive attitudes toward learning vocabulary *via* mobile phone might be influenced by its characteristics such as portability, immediacy, novelty, and legibility. With regard to the previous statement, Prensky (2005) claimed that when conducting the research regularly delivered information *via* mobile phone was manageable while the traditional manner of learning failed to arouse students' interest in study. In other words, using mobile devices as learning tools leads to enhancement of learners' motivation. The learning materials delivered by mobile devices are more appealing to students, which is not the case with paper-based learning materials. Gürkan (2018), Butarbutar (2020), and Kohnke (2020) also emphasized that using mobile apps increases students' motivation since it is more enjoyable in comparison with the traditional methods of learning. The motivation level of the participants seemed to be an important aspect of the study conducted by



Ekinci (2017) since, according to him, most of the participants consider BL as motivating and entertaining because it is considered a new method of learning. This is in line with Viberg and Grönlund (2013) who found out that the perceptions of the mobile application users are extremely positive and their motivational level is high toward mobile learning.

Another benefit of using the mobile apps in the process of vocabulary acquisition is the corrective and immediate feedback. This found concordance with Liu (2008), Klímová (2019), and Klímová and Poláková (2020). The research participants of their studies claimed that feedback helps them to learn new vocabulary items and understand their meaning which helps them to improve their language skills. The other factors affecting the learning process are the ease of use and the ubiquitous character of mobile phones. They are portable devices and can be operated anywhere and anytime with no difficulties. These factors have a stimulating effect on the students and their willingness to continue using the mobile app. These findings were supported by different studies (Gafni, 2008; Ekinci, 2017; Gafni et al., 2017; Klímová and Poláková, 2020).

However, the aim of the current study is to discover the impact of the vocabulary mobile learning application on students' range of vocabulary and to understand the students' views on the usefulness of the vocabulary mobile learning application in the process of language learning. To the best knowledge of the researchers, no study to date has investigated the overall perception of the students of the mobile applications as learning tools in EFL context in Slovakia. In order to shed more light on this subject, the authors of this study found it an opportunity to conduct research on vocabulary acquisition through the use of BL, including mobile technology, in order to discover the impact of the vocabulary mobile learning application on students' range of vocabulary, to define the usefulness of the application, and to understand students' views on the usefulness of the mobile technology in the process of language acquisition.

## MATERIALS AND METHODS

### Participants

A total number of 36 Slovak EFL students took part in this case study. The research participants were students with intermediate level of English of two different classes in the Secondary Vocational School of Gastronomy and Tourism, Nitra, Slovakia. The students were randomly assigned to an experimental group ( $N=17$ ) and a control group ( $N=19$ ). All the participants were students aged between 17 and 18 years. They had been learning English for 12 years on average. The English proficiency of the participants was measured by pre-test whose results indicated that there was no statistically significant difference between these two groups.

### Research Instrument

The mobile application *Angličtina Today*, used as the research instrument, has two separate parts. The first application part is designed as a web interface for the teacher and it is responsible

for storing information, authenticating users, efficiently collecting large data, processing, distributing messages, and responding to events. Each teacher can manage several lessons, register their students, distribute news or alerts through notifications, and respond to their comments. Using these options, the teacher can make contact with students and draw attention to the upcoming events. The key element of the web interface is the visualization of the results of all students. Based on the visualization, it is possible to evaluate each student separately, to compare the results between several study courses, or to modify the study plan (Berger et al., 2019). **Figure 1** portrays the web part of the app designed for teachers.

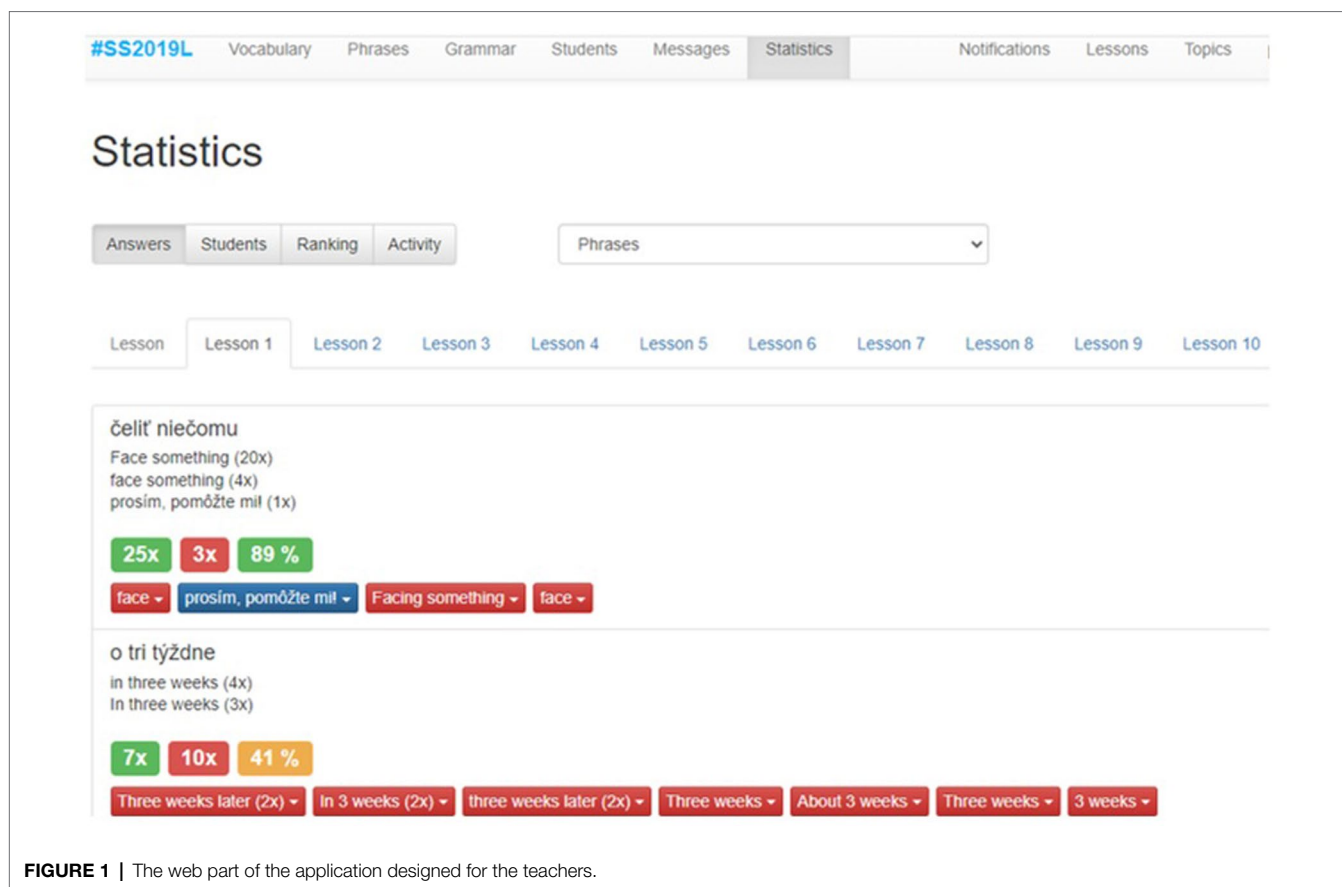
The second application part is presented with a mobile application for students. Through a mobile application, the students are enrolled into a specific course to study and test available vocabulary and phrases. For each phrase or vocabulary, students can get a translation, while using TextToSpeech technology, as well as pronunciation. The application also enables immediate communication with the teacher. At the same time, the application collects all user data and distributes it to the server part to be evaluated by teacher. The students are advised by means of notifications, e.g., to study a certain lesson (Berger et al., *ibid.*). An overview of the students' mobile app is shown in **Figure 2**.

### Procedure

This case study was conducted in the winter semester of 2021. The intervention, which lasted 10 weeks, was carried out to investigate the usefulness of the mobile application which was defined by an improved language proficiency and general overall satisfaction of the users with the application. During the first session, both experimental and control groups were given a pre-test in order to prove the same level of language proficiency. Consequently, students from the experimental group were facing BL approach. Firstly, they were introduced to new vocabulary in the traditional classroom environment; secondly, mobile application intervention was used to process the information; and thirdly, new vocabulary gained was used in context. It was done through different activities such as role-plays, discussions, or debates on the corresponding topics, in the traditional school environment. Meanwhile, students from the control group were facing traditional approach to learning. At the end of the mobile application treatment, the post-test was given to the participants of both groups. Moreover, students from the experimental group were required to answer minute papers' questions after completing each online lesson, to fill the questionnaire and participate in the virtual interview.

### Data Collection and Analysis

Based on the concept of Yin's (2009) case study, different quantitative and qualitative methods were used in order to collect and analyze data. Aiming to compare the proficiency level of the research participants before and after the mobile application intervention, standardized multiple pre-test and post-test were used in the quasi-experimental phase of the study. The test results were calculated using IBM SPSS Statistics



**FIGURE 1** | The web part of the application designed for the teachers.

26 software. After the vocabulary mobile learning application treatment, the questionnaire survey was conducted as a complementary method to quasi-experiment. The results obtained from the quantitative data collection were analyzed by descriptive statistics.

The qualitative data were collected to understand the students' perceptions of the mobile vocabulary learning application. Students were provided minute papers while using the mobile application. Despite the fact that minute papers represent a very commonly used classroom assessment technique, in the current study, it was used as the method of qualitative data collection. Moreover, virtual focus group was organized after the mobile application treatment. The ZOOM platform was used in the present study to interview the research participant. The results gained from the qualitative methods were analyzed by categorization and coding. For better understanding, the research design is provided in **Figure 3**.

## RESULTS

The aim of this study was to obtain information on the usefulness of the mobile application. On the basis of the research analyses, it may be stated that the results obtained from the quasi-experiment, questionnaire, minute papers, and virtual focus group have been in agreement. There were many results to comment on, but those which could be mutually compared

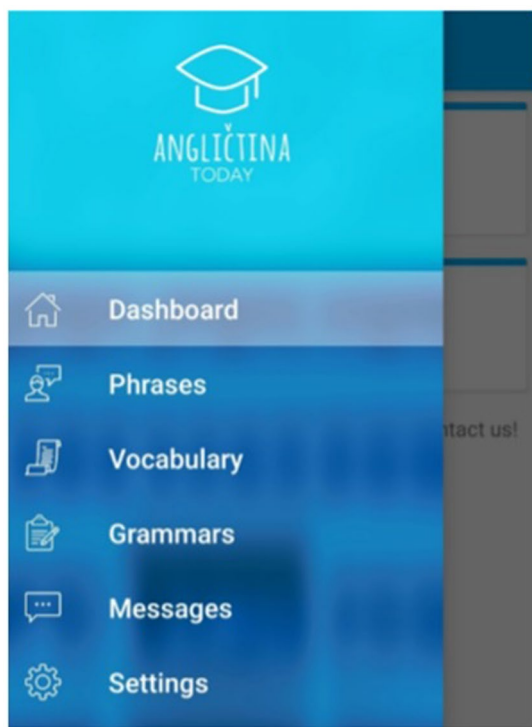
were specifically m-learners' performance, m-learners' satisfaction, m-learners' motivation, content quality, and ease of mobile application use. The comparison and the interpretation of the research results are provided below.

## Quantitative Approach to the Research

The group N consisting of 36 students was divided into two independent groups. The experimental group, the size of which was  $n_1 = 17$ , used the vocabulary mobile learning application in the process of language learning. The control group, whose size was  $n_2 = 19$ , did not use the mobile application and was taught by the use of the traditional methods. Students' results were measured by the pre-test and post-test and their values were recorded. The results were calculated using IBM SPSS Statistics 26 software. A significance level of 5% was considered when performing the tests.

In order to assess whether the students' knowledge of the subject, specifically English language, did not differ in the groups at the beginning of the study, the pre-test was revealed. The results divided according to both groups are shown in **Table 1** in which the success of students is given as percentages and **Figure 4**.

The Shapiro-Wilk test did not reject the hypothesis that both the selection in the E group (value of  $p$  0.061) and the selection in the C group (value of  $p$  0.271) came from a normal division of the students. The variances in both groups



**FIGURE 2 |** An overview of the students' application.

can be considered identical based on Levene's test ( $F=0.024$ , value of  $p$  0.877). To compare the mean values of the pre-test results, a  $t$ -test could be used for two independent samples with the same variances ( $t=-0.268$ , value of  $p$  0.790), according to which the hypothesis of equality of mean values was not rejected. It can be assumed that at the beginning of the study, i.e., at the time when the pre-test was given, the average results of students in both groups did not significantly differ. Furthermore, the following research hypothesis was stated as:

*H:* The students using the vocabulary mobile learning application achieve significantly higher learning outcomes than the students not using the application.

The box diagram of the post-test results of both groups is shown in **Figure 5**. Based on this, it can be concluded that the students from the research group (E) achieved higher results than the students from the control group (C), as expected. However, we concluded this observation by calculation. Descriptive statistics of the post-test results are provided in **Table 2**.

The Shapiro–Wilk test did not reject the hypothesis that both the selection in the experimental group (value of  $p$  0.598) and the selection in the control group (value of  $p$  0.110) came from a normal division of the students. The variances in both groups can be considered identical based on Levene's test ( $F=0.001$ , value of  $p$  0.976). To compare the mean values of the post-test results, a  $t$ -test could be used for two independent samples with the same variances ( $t=3.400$ , value of  $p$  0.002),

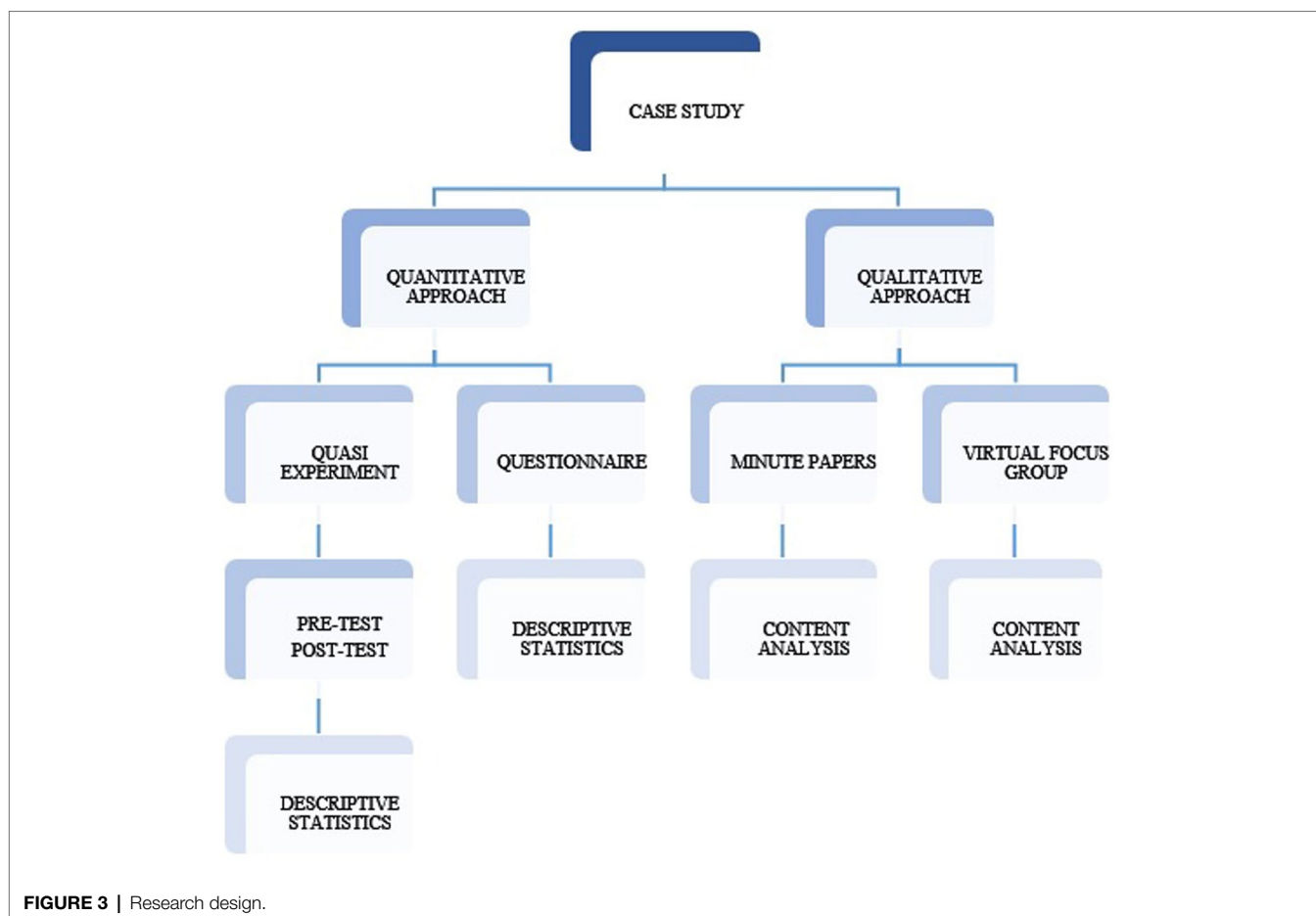
according to which the hypothesis of equality of mean values was rejected. Based on this test, it can be stated that at the level of significance of 5%, there is a statistically significant difference between the results of the post-test for the experimental and control groups. After comparison of the found estimates of the mean values of the classification in both groups, hypothesis H can be accepted.

Furthermore, the following consideration assessed whether there was a statistically significant improvement of individuals in the experimental group. Students' results in the pre-test and post-test ( $t=13.949$ , value of  $p<0.001$ ) were compared by using a paired  $t$ -test. Based on the results of this test, it can be assumed, at the level of significance of 5%, that there is a significant difference between the results of individual students in the pre-test and post-test. As the results obtained in the post-test are higher (see the line E in the **Tables 1, 2**), it can be stated that the use of a mobile application leads to a significant improvement in the learning of languages.

For a more comprehensive assessment, the ANOVA can be performed with repeated measurements using the time/test factor, whose values are pre-test and post-test, and the group factor, whose values are experimental and control groups, as two independent factors (Hendl, 2005). Thanks to the method of measurement, where the measurement was repeated on the selected research sample/students, the time/test factor can be considered to be a factor with a within-subject effect, and a group factor that divides the population can be considered to be a factor expressing an intergroup effect (between-subject effect). The Mauchly test was used to verify sphericity. ANOVA for repeated measures confirmed the effect of time/test ( $F=(1.34)=253.515$ ; value of  $p<0.001$  and correlation ratio 0.882), thus, an increase in the average score in the post-test. The group influence was not statistically significant ( $F=(1.34)=2.636$ , value of  $p$  0.114 and correlation ratio 0.072), i.e., it does not matter to which group, experimental or control, the student belongs. The calculations also show that the interaction between the group factor and the time/test factor ( $F(1.34)=51.697$ , value of  $p<0.001$  and correlation ratio 0.603) is statistically significant, which with respect to the values, given in **Tables 1, 2** means that students in the experimental group have a significantly higher average success rate in the post-test than other students. Regarding this, hypothesis H was supported.

## Qualitative Approach to the Research M-Learners' Performance

The students' perception of the vocabulary enhancement was examined by using the questionnaire and minute papers. The findings showed that a high percentage of the students (82%) perceived an improvement in vocabulary knowledge. In addition, the mobile application users (57%) stated that their vocabulary retention had been better, in other words, they remembered the vocabulary and phrases provided in the mobile app and were able to use them in the conversation, in the phase of output. Furthermore, the students stated that they had been able to express themselves better and more fluently after mobile application intervention (71%). Based on these findings, it can



**TABLE 1 |** Descriptive statistics of the experimental (E) and control (C) groups in the pre-test.

Group	Mean	Std. deviation	Median	Minimum	Maximum
E ( $n_1 = 17$ )	38.18%	16.88%	33.0%	18.00%	70.00%
C ( $n_2 = 19$ )	39.68%	16.83%	39.0%	18.00%	73.00%

be claimed that the use of the mobile application helped students to improve their language skills. Moreover, an improvement was also perceived by the research participants.

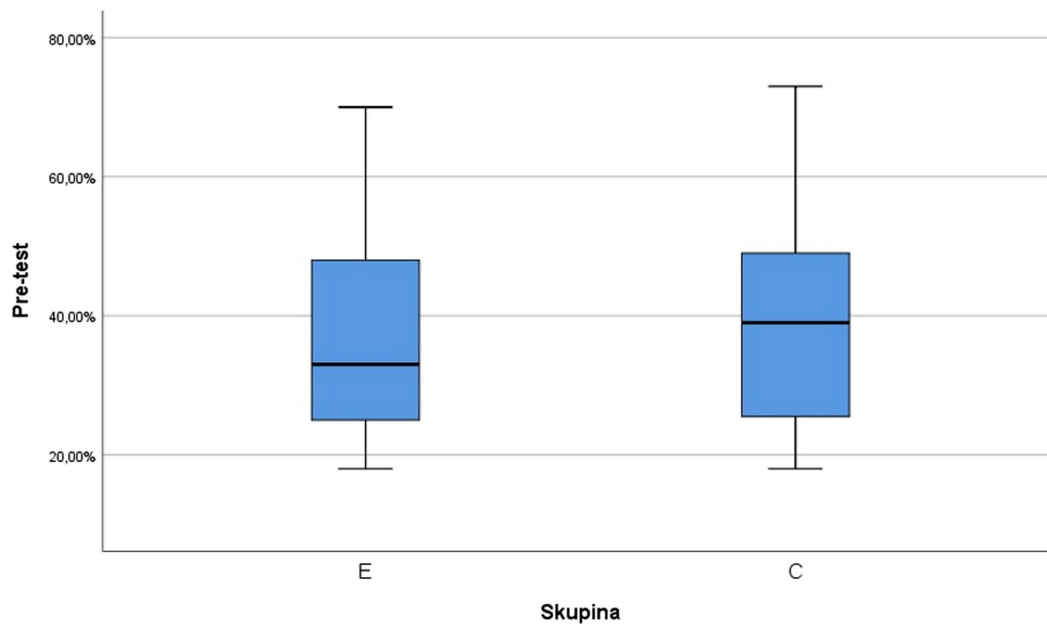
### M-Learners' Satisfaction

The questionnaire results showed that 82% of the students had been satisfied with the mobile application and would recommend it to others. This is in accordance with the minute papers and virtual focus group findings, which revealed that the students had especially appreciated an easy access to the study materials, the clear arrangement, and fast learning. Most of the students (64%) agreed that learning and practicing vocabulary through the mobile application had suited them better than learning and revising new words through the book. However, many of them also pointed out that the combination of the traditional learning and innovative methods had been

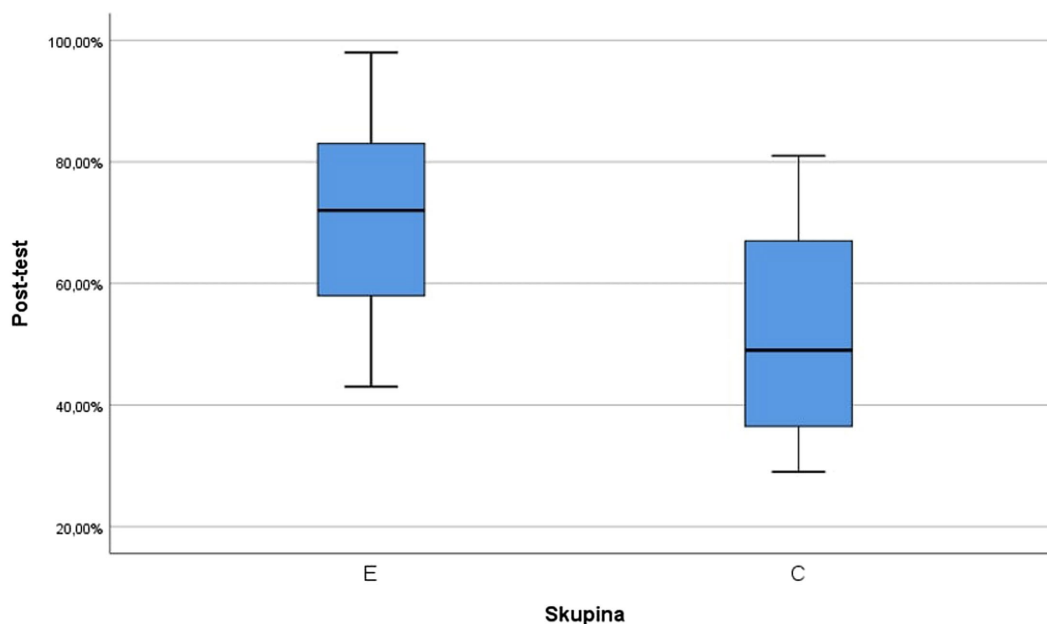
an interesting and satisfying form of learning; therefore, they (64%) would like to continue using the mobile app in the future. The results indicate the students' satisfaction with the mobile application.

### M-Learners' Motivation

The increase of the students' motivation was examined through the questionnaire, minute papers, and virtual focus group. The results revealed from the questionnaire showed that students (71%) had felt motivated by using the mobile application in the process of education. Learning was perceived as more fun and less stressful (100%). This information was also confirmed in the minute papers since many students added that their motivation was increased thanks to better learning outcomes, which were verified in the form of the additional tests. Based on the data obtained from the focus group, it can be stated that the above-mentioned tests were perceived as extrinsic motivation for students to use the mobile application. Nevertheless, such motivation was not perceived negatively by students. At the same time, mobile application users' intrinsic motivation was to increase their language skills. Regarding the findings from the various methods, the results demonstrate that the mobile application has a motivating effect on students.



**FIGURE 4 |** Box diagram of pre-test results for the experimental (E) and control (C) groups.



**FIGURE 5 |** Box diagram of post-test results for the experimental (E) and control (C) groups, 10,000%.

### Mobile Application's Functions

When comparing the data obtained from the questionnaire, the minute papers and the virtual focus group, it was found that the results of these methods matched. For a better understanding, the information provided in the mobile application was considered clear, readable, and easy to absorb (100%). The corrective feedback function was perceived as the biggest advantage of the mobile

application (100%), as it helped students to learn and understand vocabulary faster. In addition, according to the students' opinion, the corrective feedback helped them to improve and extend their vocabulary, which was also claimed by the results obtained from the post-test. On the contrary, the mobile application notifications and the pronunciation verification were considered to be the shortcomings of the mobile application (71%) since



**TABLE 2 |** Descriptive statistics of the experimental (E) and control (C) groups in the post-test.

Group	Mean	Std. deviation	Median	Minimum	Maximum
E ( $n_1 = 17$ )	71.47%	16.45%	72.0%	43.00%	98.00%
C ( $n_2 = 19$ )	52.26%	17.33%	49.0%	29.00%	81.00%

they did not work for many students during the research period. It was also confirmed that the students were not satisfied with the fact that the application could not be used offline. It can be stated that despite the benefits of the mobile application which helped students to improve their language skills, *Angličtina Today* proved to have certain shortcomings that should be eliminated in order to make it more effective in the process of language learning.

### Ease of Use

The findings revealed from the different research methods were in concordance, since they confirmed that the vocabulary mobile learning application was easy to use (100%). Students stated that using the mobile app was much easier to use in order to revise the vocabulary than using the textbooks. The mobile application users appreciated the possibility to use the mobile application anytime and anywhere, its easy manipulation, and the time-saving work with it. Therefore, it can be stated that *Angličtina Today* was perceived to be a user-friendly mobile application.

## DISCUSSION

The aim of the current research was to explore the usefulness of the mobile application defined by an improved language proficiency and also general overall satisfaction of the users with the app. The results from this study offer some key findings:

First, the questionnaire, minute papers, and virtual focus group revealed the following positive perceptions of the mobile application users:

- Vocabulary knowledge was improved after the mobile application treatment.
- Vocabulary retention was better after using the mobile application.
- Students' ability to express themselves more fluently was enhanced.
- Students would recommend *Angličtina Today* to others since the easy access, clear arrangement, and fast learning were appreciated by them.
- Learning vocabulary through the mobile application was found to be more practical than using the textbook for the same purpose.
- Students' motivation was increased thanks to the better learning outcomes.
- Learning was perceived as more fun and less stressful.
- The mobile application was considered clear, readable, and easy to absorb.

- The corrective feedback function helped students to improve their vocabulary.
- The vocabulary mobile learning application was easy to use.
- The ubiquitous character, easy manipulation, and time-saving work with the mobile application were appreciated.

These findings accord with those of Gafni et al. (2017) which revealed that using vocabulary mobile learning application has a stimulating effect on the process of learning and contributes to the willingness to continue using the application and to recommend it to others. These findings also support those of Kaceti and Klímová (2019) who indicated that the key benefits of MALL are the enhancement of the learner's cognitive capacity, the learner's motivation to study in both formal and informal settings, the learner's autonomy and confidence, as well as the promotion of personalized learning, and helping low-achieving students to reach their study goals. These findings also re-echo Nuareni's (2020) claim that the majority of the students has positive perception on the usage of MALL in learning English language. These findings are also in line with the most recent study by Jeong (2022), who claim that MALL helped his students with access to learning contents, created flexible and self-directed learning environment, better interaction, and improved self-efficacy in English learning performance.

Second, the implementation of the vocabulary mobile learning application was effective because the students from the experimental group reached better results than the students from the control group. Moreover, an enhancement of the language skills was also perceived by the mobile application users. This is in line with the results of several other studies that implemented MALL in the process of education when conducting research. For instance, Alqahtani and Mohhammad (2015) claimed that using the mobile learning application enhanced students' performance as well as their satisfaction. Similar results were found by Lu (2008) who stated that the participants from the experimental group identified more vocabulary than the group learning vocabulary through the printed materials. The same findings were revealed by Palalas (2011) and Gharehblagh and Nasri (2020) who indicated that MALL helped to improve the vocabulary knowledge of EFL students.

Overall, the results obtained from the different quantitative and qualitative research methods proved a high level of the mobile application usefulness. This claim is based on the fact that using mobile application in the process of education helped students to reach better results in the post-test, helped to improve students' vocabulary and the ability to express themselves more fluently, proved to be user-friendly, and enhanced students' motivation to learn. Nevertheless, as Klimova (2021) states, teachers have to guide their students in the use of such an app to make the learning process effective and meaningful.

This study helps to understand the perceived advantages and disadvantages of using vocabulary mobile learning application in blended foreign language learning. It turned out that most of the participants found the mobile application

beneficial additional learning tool since it helped them to improve their language skills and made learning more fun and less stressful. The obtained research findings can help foreign language teachers to understand the benefits of combining mobile learning with traditional face-to-face learning. It can motivate them to innovate their teaching and lose the fear of using technology due to its seemingly disruptive nature. Furthermore, it might motivate foreign language students to use their mobile devices as the learning tool. Moreover, the findings can assist developers of vocabulary mobile learning applications since students' preferences are provided in the present study.

## IMPLICATIONS

In view of the results of this study, some pedagogical recommendations are suggested to EFL instructors. Teachers should be the role models for their students and **teach them how to use mobile phones in the process of language learning**. Despite the fact that students use their mobile phones on a daily basis, they might not know how to use them as learning tools. Therefore, they need the guidance of a teacher.

It is sufficient for MALL to be implemented only in one of the learning phases, preferably during the information processing phase. Students can be introduced new vocabulary in the classroom environment through the traditional methods of learning and, subsequently, **the new vocabulary can be revised through a mobile application, outside the school environment**. The last phase of output should take place in the classroom, where the various activities can be used to verify students' vocabulary knowledge.

When revealing the mobile lessons, it is important to provide learning materials that are in concordance with the curriculum taught in the classroom. **The information provided should be well organized, readable, and easy to absorb.**

It is not necessary to provide a large amount of vocabulary in each mobile lesson. However, **it is recommended to add new vocabulary regularly**, preferably after each face-to-face lesson.

Teachers should regularly **send notifications** and remind their students of new material or testing. If mobile application notifications do not work, it is necessary to look for new solutions to inform students about these events.

**It is important to carry out a formative assessment**, which makes it possible to monitor students' progress. The results obtained from the assessment can help teachers to understand the needs of the students and possibly change the learning plans. Formative assessment can also be a motivation for students to use the mobile application; moreover, students are provided with feedback on their learning performance.

It is essential for teachers to **ask for feedback from the students** and understand their perception of the use of new technologies in the learning process. It can help teachers

to improve their teaching and, what is more, students value being asked for their opinion since they feel being recognized when they are given space to express their views. Such an approach can help teachers to create a positive school climate and thus improve the learning process.

It is recommended to **give students a free choice** and let them decide whether they want to use the mobile technologies in the process of learning or prefer traditional learning through textbooks. Forcing students to use the mobile technologies could provoke resistance to learning. Students must not be punished for their decisions.

Furthermore, it would be suitable to **raise other teachers' awareness** of the implementation of the mobile technologies in the process of the language learning which can help students to improve their language skills. The older generation of teachers often resist the use of innovative trends, but nowadays, the implementation of the technology in education seems to be essential. The COVID-19 pandemic clarified the importance of the technologies as traditional learning suddenly switched into distance learning. This could not have worked without the inclusion of the technology in the process of learning.

## LIMITATIONS

Naturally, there are certain limitations to be considered in this study. To start with, the research sample size was small; therefore, it was difficult to get more statistically significant results and generalize the study. Additionally, the time the secondary school students experienced the mobile application was relatively short. A longer period may have provided an additional and deeper insight. Furthermore, there was no space to examine the vocabulary retention perceived by the students in the follow-up period. However, this fact was confirmed by different studies whose results were in line with the findings of the present research.

## DATA AVAILABILITY STATEMENT

The raw data supporting the conclusions of this article will be made available by the authors, without undue reservation.

## AUTHOR CONTRIBUTIONS

All authors listed have made a substantial, direct, and intellectual contribution to the work and approved it for publication.

## FUNDING

The paper is supported by the project Excellence (2202/2022) at the Faculty of Informatics and Management, University of Hradec Králové, Czechia.

## REFERENCES

- Alqahtani, M. (2015). The importance of vocabulary in language learning and how to be taught. *Int. J. Teach. Educ.* 3, 21–34. doi: 10.20472/TE.2015.3.3.002
- Alqahtani, M., and Mohamad, H. (2015). Mobile applications' impact on students' performance and satisfaction. *J. Educ. Technol.* 14, 102–112. doi: 10.29252/ijtee.3.2.48
- Berger, A., Klímová, B., and Poulová, P. (2019). Mobile application as support of English on-line learning. *Educ. Comput. Sci.* 647–653. doi: 10.5220/0007732506470653
- Birkner, V. A. (2016). Revisting input and output hypotheses in second language learning. *Asian Educ. Stud.* 1:19. doi: 10.20849/aes.v1i1.18
- Björqvinnsson, L. A. (2012). Incidental and instructional vocabulary. Available at: [https://www.academia.edu/3321504/Incidental\\_and\\_Instructional\\_Vocabulary\\_Acquisition](https://www.academia.edu/3321504/Incidental_and_Instructional_Vocabulary_Acquisition) (Accessed January 3, 2022).
- Butarbutar, R. (2020). "Mobile assisted language learning," in *Advances in Social Science, Education and Humanities Research, Proceedings of the 3rd International Conference on Social Sciences*. Atlantis Press, 390–392.
- Council of Europe (2018). Available at: <https://www.coe.int/en/web/common-european-framework-reference-languages> (Accessed January 19, 2022).
- Ekinci, M. (2017). Perceptions of EFL learners about using mobile applications for English language learnings: a case study. *Int. J. Acad. Res.* 5, 175–193. doi: 10.18033/ijla.3659
- Eraut, M. (2004). The practice of reflection. *Learn. Health Soc. Care* 3, 47–52. doi: 10.1111/j.1473-6861.2004.00066.x
- Gafni, R. (2008). *Measuring Quality of M-Learning Systems*. Santa Rosa, CA: Informing Science Press.
- Gafni, R., Achituv, D. B., and Rachmani, G. J. (2017). Learning foreign languages using mobile applications. *J. Inf. Technol. Educ. Res.* 13, 49–89.
- Gürkan, S. (2018). The effects of a mobile assisted vocabulary learning application on vocabulary learning. *Turkish J. Qual. Inq.* 9, 288–311. doi: 10.17569/tojqi.407512
- Gharebblagh, N. M., and Nasri, N. (2020). Developing EFL elementary learners' writing skills through mobile-assisted language learning (MALL). *Teach. Engl. Technol.* 20, 104–121.
- Hendl, J. (2005). *Kvalitativní výzkum: Základní metody a aplikace*. Praha: Portál.
- Jeong, K. O. (2022). Facilitating sustainable self-directed learning experience with the use of mobile-assisted language learning. *Sustainability* 14:2894. doi: 10.3390/su14052894
- Kacetyl, J., and Klímová, B. (2019). Use of smartphone applications in English language learning: a challenge for foreign language education. *Educ. Sci.* 9:179. doi: 10.3390/educsci9030179
- Klímová, B. (2017). Mobile phones and/or smartphones and their apps for teaching English as a foreign language. *Educ. Inf. Technol.* 23, 1091–1099. doi: 10.1007/s10639-017-9655-5
- Klímová, B. (2019). Impact of mobile learning on students' achievement results. *Educ. Sci.* 9:90. doi: 10.3390/educsci9020090
- Klimova, B. (2021). Evaluating impact of mobile applications on EFL university learners' vocabulary learning: a review study. *Procedia Comput. Sci.* 184, 859–864. doi: 10.1016/j.procs.2021.03.108
- Klímová, B., and Poláková, P. (2020). Students' perception of an EFL vocabulary learning mobile application: a case study. *Educ. Sci.* 10:37. doi: 10.3390/educsci10020037
- Kohnke, L. (2020). Exploring learner perception, experience and motivation of using a mobile app in L2 vocabulary acquisition. *IJCALLT* 10, 15–26. doi: 10.4018/IJCALLT.2020010102
- Kukulka-Hulme, A. (2009). Will mobile learning change language learning? *ReCALL* 21, 157–165. doi: 10.1017/S0958344009000202
- Liu, M. (2008). Effectiveness of vocabulary learning via mobile phone. *J. Comput. Assist. Learn.* 24, 515–525. doi: 10.1111/j.1365-2729.2008.00289.x
- Lu, M. (2008). Effectiveness of vocabulary learning via mobile phone. *J. Comput. Assist. Learn.* 24, 515–525. doi: 10.1111/j.1365-2729.2008.00289.x
- Mahdi, S. H. (2017). Effectiveness of mobile devices on vocabulary learning: A meta-analysis. *J. Educ. Comput. Res.* 56, 134–154. doi: 10.1177/0735633117698826
- Mikulecký, P. (2019). "Blended learning in smart learning environments," in *EPIA Conference on Artificial Intelligence*. August 24–27, 2020.
- Nasab, H. S., and Taki, S. (2016). Effects of MALL in blended learning on Iranian EFL learners' reading comprehension. *Modern J. Lang. Teach. Methods.* 6, 854–869.
- Nation, I. S. P. (2001). *Learning Vocabulary in Another Language*. Cambridge, England: Cambridge University Press.
- Neumeier, P. (2005). A closer look at blended learning: parameters for designing a blended learning environment for language teaching and learning. *ReCALL* 17, 163–178. doi: 10.1017/S0958344005000224
- Nuareni, C. (2020). Mobile-assisted language learning (MALL): Students' perception and problems towards Mobile learning in English language. *J. Phys. Conf. Ser.* 1641:7.
- Palalas, A. (2011). "Mobile-Assisted Language Learning: Designing for Your Students," in *Second Language Teaching and Learning with Technology: Views of Emergent Researcher*. Dublin: Research-Publishing.net.
- Poláková, P. (2019). Using mobile technologies in foreign Language Learning. *Zborník štúdií z medzinárodnej vedeckej konferencie Inovatívne trendy v odborových didaktikách. Prepojenie teórie a praxe výučbových stratégií kritickeho a tvorivého myslenia*.
- Prensky, M. (2005). What can you learn from a cell phone? Almost anything! Available at: <https://archive.cecis.ori.hr/app/public/conferences/1/cecis2015/papers/745.pdf> (Accessed December 20, 2021).
- Rezaei, A., Mai, N. B., and Pesaranhader, A. J. (2013). "Effectiveness of using english vocabulary mobile applications on ESLs learning performance," in *International Conference on Informatics and Creative Multimedia, 2013*. Kuala Lumpur, Malaysia; IEEE; 114–118. September 4–6, 2013
- Schmitt, N. (2010). *Researching Vocabulary: A Vocabulary Research Manual*. Great Britain: Palgrave Macmillan.
- Senffner, D., and Kepler, L. G. (2015). *Blended Learning that Works*. Alexandria, VA: Association for Talent Development.
- Soleimani, H., Mohammaddokht, F., and Fathi, J. (2022). Exploring the effect of assisted repeated Reading on incidental vocabulary learning and vocabulary learning self-efficacy in an EFL context. *Front. Psychol.* 13:851812. doi: 10.3389/fpsyg.2022.851812
- Viberg, O., and Grönlund, A. (2013). Cross-cultural analysis of users' attitudes toward the use of mobile devices in second and foreign language learning in higher education: a case from Sweden and China. *Comput. Educ.* 69, 169–180. doi: 10.1016/j.compedu.2013.07.014
- Wilkins, D. (1972). *Linguistics in Language Teaching*. London, England: Arnold.
- Yin, R. K. (2009). *Case Study Research: Design and Methods. 4th Edn*. Thousand Oaks, CA: Sage publications.
- Zhang, (2009). The role of input, interaction and output in the development of oral Fluency. *CCSE J.* 2.
- Zhang, H. (2011). Reexamining the effectiveness of vocabulary learning via mobile phones. *TOJET* 10, 203–214.
- Zhang, D., and Feng Teng, M. (2012). The associations between working memory and the effects of multimedia input on L2 vocabulary learning. *Int. Rev. Appl. Ling. Lang. Teach.*

**Conflict of Interest:** The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

**Publisher's Note:** All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Copyright © 2022 Polakova and Klimova. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.





# On-Screen Texts in Audiovisual Input for L2 Vocabulary Learning: A Review

Rong Wei<sup>1</sup> and Lin Fan<sup>2,3\*</sup>

<sup>1</sup>School of Foreign Languages, Ma'an Shan University, Ma'an Shan, China, <sup>2</sup>National Research Center for Foreign Language Education, Beijing Foreign Studies University, Beijing, China, <sup>3</sup>School of Foreign Languages, Qingdao University of Science and Technology, Qingdao, China

## OPEN ACCESS

### Edited by:

Ehsan Rassaei,  
Majan University College, Oman

### Reviewed by:

Musa Nushi,  
Shahid Beheshti University, Iran  
Leila Gholami,  
Arizona State University,  
United States  
Nasim Ghanbari,  
Persian Gulf University, Iran

### \*Correspondence:

Lin Fan  
fanlinqd@163.com

### Specialty section:

This article was submitted to  
Language Sciences,  
a section of the journal  
Frontiers in Psychology

**Received:** 25 March 2022

**Accepted:** 25 April 2022

**Published:** 13 May 2022

### Citation:

Wei R and Fan L (2022) On-Screen  
Texts in Audiovisual Input for L2  
Vocabulary Learning: A Review.  
Front. Psychol. 13:904523.  
doi: 10.3389/fpsyg.2022.904523

Audiovisual input has received increasing attention from the Second Language Acquisition (SLA) and the Computer-Assisted Language Learning (CALL) domains during the past few decades due to its vividness, authenticity, and easy accessibility. Videos with on-screen texts, as a widespread way of audiovisual input in second language (L2) teaching and learning, influence L2 learners' performance in various aspects, including their vocabulary learning. The wide application and profound influence of such kind of input call for a systemic review on this important domain of research. Accordingly, this paper reviews the empirical studies on the effects of on-screen texts on L2 vocabulary learning. Specifically, it seeks to evaluate the role of different types of on-screen texts (i.e., subtitles, captions, and dual subtitles) and various modes of captions (i.e., full captions, keyword captions, glossed captions, annotated captions, and enhanced captions) in L2 vocabulary development. It also discusses other factors that concur with on-screen texts and influence L2 vocabulary gains from audiovisual input, such as learners' vocabulary size, L2 proficiency, frequency of occurrence, number of viewing, instructional strategy, and test time. Finally, some suggestions are provided for future research.

**Keywords:** audiovisual input, on-screen text, subtitle, caption, L2 vocabulary, CALL

## INTRODUCTION

Audiovisual input,<sup>1</sup> whose application speeds up owing to the development of multimedia technology, has received growing attention from the SLA and the CALL domains during the past few decades. It not only provides vivid and authentic language learning materials, but also expands the way we access new information, even optimizes the use of our cognitive resources. Specifically, compared with traditional written input (i.e., textbooks), audiovisual input offers contextual learning environment which demonstrates language in use vividly and authentically; it is no longer limited to static information from printed words, but provides easy access to updated dynamic materials through TV, computers, and mobile devices; it conveys information that fully activates learners' visual and auditory systems in cognition so as to optimize the learning outcome (Low and Sweller, 2014).

Audiovisual input contains several key elements, including audio, video, and on-screen text. In the fields of SLA and CALL, on-screen text is a general term for the text about the video content displayed synchronously or asynchronously on the screen (Mohsen, 2016a). It usually

<sup>1</sup>"Audiovisual input" and "videos with on-screen texts" are used interchangeably in this review.

presents itself in different types—subtitle (L1 text, L2 video), caption (L2 text, L2 video), reversed subtitle<sup>2</sup> (L2 text, L1 video) and dual subtitle (L1 + L2 texts, L2 video), or in different modes (in the case of captions)—keyword caption (L2 keywords, L2 video), glossed caption (L2 text + L1 gloss, L2 video), and so forth (e.g., Sydorenko, 2010; Lwo and Lin, 2012; Montero Perez et al., 2013, 2015, 2018; Fievez et al., 2021).

As a widespread way of audiovisual input in SLA and CALL, videos with on-screen texts have influenced L2 learners' performance from all-round aspects, including pronunciation (e.g., Mitterer and McQueen, 2009; Wisniewska and Mora, 2020; Wong et al., 2020; Mohsen and Mahdi, 2021), grammar (e.g., Lee and Révész, 2018, 2020; Pattemore and Muñoz, 2020), writing (e.g., Alobaid, 2021), pragmatics (e.g., Barón and Celaya, 2022), listening comprehension (e.g., Montero Perez et al., 2014a; Liao et al., 2020; Pujadas and Muñoz, 2020; Lee et al., 2021), as well as vocabulary learning<sup>3</sup> (e.g., Montero Perez et al., 2014b; Peters et al., 2016; Teng, 2020; Muñoz et al., 2021). The general positive role of on-screen texts was reported in these aspects. In particular, empirical research on the role of on-screen texts in L2 vocabulary learning have yielded fruitful but inconclusive results when they compared the effects of on-screen texts in various forms, which is the focus of the present review. Generally, some researchers have explored the effects of the two major types of on-screen texts and observed that captions exerted greater influence than subtitles in promoting word learning (e.g., Frumuselu et al., 2015; Peters et al., 2016; Peters, 2019; Baranowska, 2020), while others claimed that types of on-screen texts had no significant correlations with learners' performance on vocabulary growth (e.g., Lwo and Lin, 2012; Muñoz et al., 2021). Still another group concentrating on the diversified caption modes (e.g., keyword caption, glossed caption, annotated caption, and enhanced caption) obtained even more multifarious results.

Given that researchers have yielded fruitful results of the role of on-screen texts in L2 learning, reviews were conducted to depict the research status and development in this field. Vanderplank (2010), one of the pioneers, gave an assessment of primary research studies on language teaching and learning *via* television and video and highlighted the role of subtitles and captions. Thereafter, reviews on L2 learning aided by videos with on-screen texts have emerged intermittently. Some of them were comprehensive ones which discussed multiple aspects of L2 learning (e.g., Matielo et al., 2015), while others focused on one or two specific aspects, among which listening comprehension (e.g., Montero Perez et al., 2013; Mohsen, 2016a; Yeldham, 2018) and vocabulary learning (e.g., Montero Perez et al., 2013; Mohsen, 2016a) were the most popular topics in these reviews. Nonetheless, these reviews barely offer an in-depth description about the field of on-screen-text-aided

L2 learning mainly in consequence of the inadequate empirical research. Fortunately, this field has met its heyday thanks to the ever-developing multimedia technology. Empirical studies have sprouted in the past few years, which call for latest reviews on literature. Most recently, Montero Perez (2022) presents a panorama of the status quo by examining a variety of documents ranging from books and edited volumes to reviews, conferences, special issues, as well as empirical studies, and encompassing such inclusive L2 learning aspects as comprehension, vocabulary, grammar and listening. The existing reviews have undoubtedly made great contributions to the academic landscape of CALL or SLA supported by videos with on-screen texts. Since vocabulary plays a predominant role in L2 learning, as indicated in a much-cited statement that “Without grammar very little can be conveyed; but without vocabulary nothing can be conveyed” (Wilkins, 1972, p. 111), “most [empirical] studies have looked into the potential of on-screen text for vocabulary learning” (Montero Perez, 2022, p. 20). Yet there is few review specifically taking vocabulary learning as the sole object to delve into the effects of on-screen texts. To our knowledge, Teng (2021) was the only one who paid exclusive attention to incidental vocabulary acquisition from captioned videos, but only focused on 6 studies in his review.

Hence, it is urgent to launch a review to sort out the literature on the effects of on-screen texts on L2 vocabulary learning. To identify the scope of this review, several criteria were used to select the literature. First, research articles, monographs, and book chapters that employ experimental designs are included. Doctoral dissertations, conference proceedings, editorial materials, book reviews are all excluded. Second, since the application of audiovisual input to L2 learning has been greatly accelerated *via* the flourishing Internet-related platforms during the past decade, that is, from 2012 to 2022, this review mainly concentrates on but is not confined to this time span. Prominent earlier studies would also be mentioned to demonstrate their contributions to this area. Third, studies published in languages other than English are ignored. Fourth, the research area is restricted to SLA and CALL. Articles discussing the role of on-screen texts in supporting pre-school children or the deaf or hard-of-hearing, or studies addressing other issues (e.g., reading behavior or working memory) with on-screen texts merely as the context are all expunged. Last but not least, videos with on-screen texts in this review are defined as a kind of audiovisual input which is composed of dynamic visuals instead of static pictures or graphics, audio with L2 soundtrack, and different forms of on-screen texts.

Following the criteria of data retrieval, this review sets out to paint a comprehensive picture about the effects of videos with on-screen texts on L2 vocabulary learning. It begins with the theories which lay a solid foundation for research on audiovisual input in L2 learning. Then, based on the clarification of terminology and categories of on-screen texts in SLA and CALL contexts, the retrieved empirical studies are reviewed mainly in two hierarchical groups. Meanwhile, some learner-related and experiment-related factors that concur with on-screen texts to influence L2 vocabulary development are also appraised.

<sup>2</sup>Since this review focuses on vocabulary learning in L2 context, reversed subtitle which accompanies L1 video is ignored.

<sup>3</sup>“Vocabulary learning” and “vocabulary acquisition” are used interchangeably in this review, though minor difference exists between them, that is, the former usually stresses an intentional learning behavior while the latter refers to an incidental by-product of learning something else.

Finally, suggestions for future research on on-screen-text-aided L2 vocabulary learning are provided.

## RATIONALE FOR AUDIOVISUAL INPUT IN L2 LEARNING

### Dual Coding Theory

The origin of dual coding theory (DCT) can be traced back to Paivio (1969) review on the studies of imagery. Paivio (1990) believed that humans are unique in nature in that they have the innate ability to deal concurrently with modality-specific verbal and non-verbal representations which herein refer to language and imagery, respectively. Accordingly, the fundamental assumption in DCT is that human cognition in reading and writing depends on dual coding systems of mental representations. The one in charge of the verbal representations is defined as the verbal system specialized for language; the other in charge of the non-verbal representations is the imagery system specialized for non-verbal objects and events (Paivio, 1990, 2010; Sadoski and Paivio, 2000). The two systems function independently but interconnectedly, that is, they can be activated individually or simultaneously, depending on the nature of the task. Both systems contribute to human cognition in reading and writing, the operating units of which are, respectively, called logogen and imagen (Paivio, 2010).

The idea of dual coding systems in handling environmental information, then, enlightened research on on-screen-text-aided L2 learning from audiovisual input. The visual information in videos forms imagens and the soundtrack along with the on-screen texts constitutes corresponding logogens, which together contribute to the process of new information in cognition. Since human cognition functions in two systems according to this theory, the combination of visual and auditory stimuli should outperform any single stimulus, thus laying a foundation for the merits of audiovisual input.

### Cognitive Load Theory

Cognitive load theory (CLT) was put forward by Sweller in 1988 based on human cognitive architecture. According to CLT, cognitive architecture is composed of “working memory, long-term memory and the relationships between them” (Sweller et al., 2019, p. 263). The process of learning is largely confined to our working memory which can only handle a limited number of information elements for a limited duration (*ibid.*). Hence, the more requirements a learning task imposes, the higher the cognitive load is. When it is beyond the capacity of the working memory, it hinders information transformation and knowledge construction into the long-term memory (*ibid.*). On the other hand, if the cognitive resources are reasonably allocated to information from different modalities (i.e., visual and auditory), cognitive load will be reduced and learning will be more effective than from a single modality (Low and Sweller, 2014). As a result, the modality effect, one of the cognitive load effects, also lends support to audiovisual input for L2 learning.

## Multimedia Learning Theory

Multimedia learning theory proposed by Mayer is closely associated with the above two theories. It hypothesizes that “people can learn more deeply from words and pictures than from words alone” (Mayer, 2014, p. 1). This theory entails three assumptions: the dual-channel assumption (i.e., there are two channels for processing visual and auditory information separately), the limited capacity assumption (i.e., each channel has a limited capacity), and the active processing assumption (i.e., humans are active processors of the ongoing information; *ibid.*). It provides an explanation to the cognitive processes of language learning from audiovisual input: learners first select useful words and images, then organize them into coherent verbal and pictorial representations, and finally integrate these representations with their existing knowledge (*ibid.*). In this theory, multimedia instruction principles inspired by cognitive load effects (e.g., the split-attention effect, the redundancy effect, and the modality effect) were elaborated and thereafter have shown far-reaching significance in explaining the various outcomes of empirical research on L2 learning from audiovisual input.

## CATEGORIES OF ON-SCREEN TEXTS

Even though audiovisual input has been applied to L2 teaching and learning for a few decades, the divergence in using terminology has not been settled yet. Some studies did not make a distinction between “subtitle” and “caption.” Instead, they commonly employed “L1/L2 subtitles” (e.g., Bisson et al., 2014; Vulchanova et al., 2015; Baranowska, 2020; Hao et al., 2021), or occasionally “L1/L2 captions” (e.g., Lwo and Lin, 2012) to refer to interlingual and intralingual on-screen texts, respectively. However, this phenomenon may lead to confusion for readers and novice researchers.

To clarify these two concepts, we need to summarize the history of their applications first. Initially, subtitle was one of the favorite adaptation methods in European countries in the 1990s when people were frequently introduced to foreign language TV programs (Koolstra et al., 2002). It presented (condensed) translation in the viewers’ native language along with the original foreign language soundtrack (Peters et al., 2016). Gradually, its value in fostering informal language learning was detected (*ibid.*). By contrast, caption was, originally, not designed to oblige the common people, but rather the deaf or hard-of-hearing (Vanderplank, 2010; Bisson et al., 2014; Peters et al., 2016; Teng, 2021). Like subtitle, caption was then regarded as a scaffold to facilitate L2 learning provided that the input was not too challenging for the learners’ language proficiency (Danan, 2004; Vanderplank, 2010). While the effect of subtitles on L2 learning was widely scrutinized in earlier research carried out in the 1990s, caption has been under the spotlight as a more recent concern (Peters et al., 2016). Consequently, various terms, derived from different caption modes, were engendered to meet their corresponding research purposes.

This review tentatively categorizes the empirical studies into two groups according to the hierarchical order of their involved on-screen texts. The one on the upper level pertaining to different

types of on-screen texts (i.e., subtitle, caption, and dual subtitle) concerns the optimum language(s) to be displayed with videos; and the other on the lower level of different modes of captions (i.e., full caption, keyword caption, glossed caption, annotated caption, and enhanced caption) is to further investigate the favorable mode of displaying one particular type of on-screen text—caption, which is generally supposed to be beneficial. In SLA and CALL contexts, we adopt the definitions prevailing in most studies. Subtitle is the interlingual on-screen text which provides L1 translation to the L2 soundtrack, and caption is the intralingual on-screen text which provides L2 verbatim transcription to the L2 soundtrack (Danan, 2004; Winke et al., 2010; Hsu et al., 2013). Dual subtitle is the one that combines L1 translation and L2 verbatim transcription simultaneously (Lwo and Lin, 2012; Hao et al., 2021; Wang and Pellicer-Sánchez, 2022). Full caption, another term for caption, is employed when it is discussed in the scope of caption modes to differ from keyword caption in particular. Keyword caption is the caption mode which only contains one or a maximum of three or four consecutive words that are crucial for constructing sentence meaning (Montero Perez et al., 2018; Teng, 2019). Glossed (full) caption is defined as complete captions in which some keywords include access to contextual meaning in viewers' native language (Teng, 2020). Annotated caption is a kind of full caption with some keywords connected to an annotation which contains the L1 and/or L2 definition(s) of the word, an L2 example sentence, an image for illustration, and sometimes the words' pronunciation (Aldera and Mohsen, 2013; Mohsen, 2016b). Enhanced caption is a kind of full caption with some keywords bolded, underlined, or colored to enhance their salience (Cintrón-Valentín and García-Amaya, 2021; Majuddin et al., 2021). **Figure 1** shows the main categories of on-screen texts and their relations.

## EFFECTS OF ON-SCREEN TEXTS ON L2 VOCABULARY LEARNING

It is acknowledged that on-screen texts do exert a positive influence on L2 vocabulary learning. Yet researchers have never stopped moving forward to explore the most beneficial language to be displayed and the most appropriate way of

displaying the language on the screen. In line with the foci of the empirical studies, we attempt to answer the above two questions by investigating, respectively, the effects of different types of on-screen texts and the effects of different modes of captions. Since word knowledge is a multi-dimensional concept that involves form, meaning, use, and each of them can be further divided into more detailed branches (for details, see, e.g., Teng, 2021), the retrieved empirical studies may explore one or several aspects of word knowledge so as to contribute to the whole view of on-screen-text-aided L2 vocabulary learning.

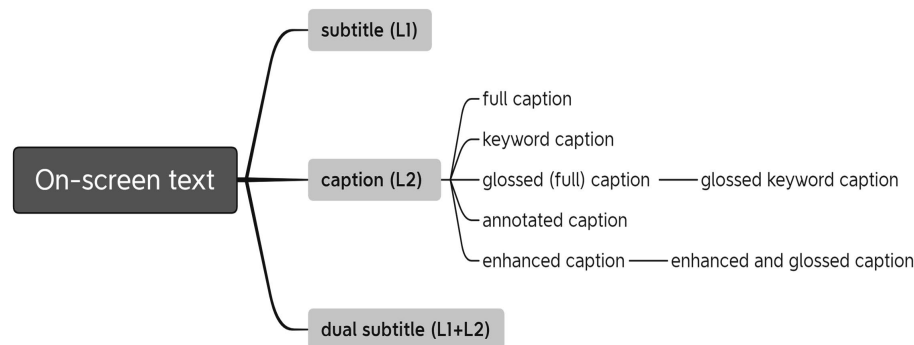
## Effects of Different Types of On-Screen Texts

The positive effects of subtitles (e.g., Koolstra and Beentjes, 1999) and captions (e.g., Neuman and Koskinen, 1992; Yuksel and Tanriverdi, 2009; Sydorenko, 2010; Winke et al., 2010) on vocabulary growth were identified separately in earlier studies. Recently, researchers tend to figure out which type of on-screen text (i.e., language to be displayed on the screen) is the optimum for L2 vocabulary gains. In addressing this issue, three camps were formed according to their research results (**Table 1**).

The first camp held that captions were superior to subtitles. Peters et al. (2016) carried out two experiments, respectively, on intermediate and low-proficiency English-as-a-foreign language (EFL) students to investigate the differential effects of subtitles

**TABLE 1** | Three camps in addressing the optimum type of on-screen text.

	Synchronic studies	Longitudinal studies
Camp 1 (Captions were better)	Peters et al., 2016 Peters, 2019 Baranowska, 2020 Wang and Pellicer-Sánchez, 2022	Frumuselu et al., 2015
Camp 2 (Irrelevant to types of on-screen texts)	Lwo and Lin, 2012 Bisson et al., 2014 Vulchanova et al., 2015 Birulés-Muntané and Soto-Faraco, 2016	Frumeselu, 2019 Muñoz et al., 2021
Camp 3 (Subtitles were better)		Hao et al., 2021



**FIGURE 1** | Categories of on-screen texts and their relations.



and captions. The two experiments almost arrived at the same conclusion that captions showed greater influence on word form than subtitles. However, the results may be partially explained by the fact that participants' unfamiliarity with captions aroused their additional attention to the form of the displayed target words. Despite the small sample size and the general low learning gains in all groups, this study provided enlightening directions for future research, such as the combination of captions and imagery, which illuminated a subsequent experiment of great importance in this field (i.e., Peters, 2019). Peters (2019) and Baranowska (2020) which invited intermediate EFL students as their participants both showed that the caption groups performed better than the subtitle groups and the control groups without any on-screen text. A more recent study conducted by Wang and Pellicer-Sánchez (2022) supplemented a less discussed type—dual subtitle—while comparing the effects of different on-screen text types. They found that captions significantly outperformed the other three types (i.e., dual subtitle, subtitle, and no subtitle) in form recognition, and dual subtitles showed greater potential in promoting meaning recall and meaning recognition. This eye-tracking study not only demonstrated the positive role of dual subtitles on meaning knowledge, but also confirmed participants' preference for L1 translation to L2 form in dual subtitles with online evidence. A longitudinal study (i.e., Frumuselu et al., 2015) tested the informal and conversational speech learning outcome of university undergraduates with different L2 proficiency from A2-C1 of CEFR after a 7-week intervention. The results again confirmed the advantage of captions, independent on L2 proficiency.

The findings of the second group showed that L2 vocabulary learning had little to do with types of on-screen texts. Lwo and Lin (2012) were among the few researchers who introduced dual subtitles into their study, yet their result was totally different from that in Wang and Pellicer-Sánchez (2022). Instead of observing the positive impact of captions or dual subtitles, they found that neither the existence nor the types of on-screen texts exerted any influence on junior high school students' vocabulary recognition and use, which was ascribed to the excessive visual and auditory support that dwarfed the effects of on-screen texts in the teaching material. In other words, the audiovisual input was so easy that the learners did not rely on the on-screen texts anymore. Contrary to the easy input, more studies showed no vocabulary gains from audiovisual input irrespective of the on-screen text types under the circumstances that the participants were completely new to the target L2 (Bisson et al., 2014) and that the input full of low frequency words was far beyond the learners' linguistic ability (Birulés-Muntané and Soto-Faraco, 2016). These may indicate that the on-screen texts would turn out to be in vain when the input is either too easy or too difficult. So the above synchronic results may be untenable as a consequence of the inappropriateness of learning materials. Vulchanova et al. (2015) compared the long-term effect of subtitles and captions by measuring learners' word definition and word recall performance 4 weeks after viewing the learning material. Although an age-modulated effect was noticed in the subtitle condition in word definition task, no effects of on-screen text types were observed in separate age group analyses in either

word definition or word recall. The discouraging result was ascribed to two reasons: the long interval between the viewing and the tests, and the one-off intervention which provided no more encounters. In addition, it was worth mentioning that the research of Frumuselu (2019), almost identical to that of Frumuselu et al. (2015) in experimental design, reported that neither subtitles nor captions stimulated the learning of informal and colloquial vocabulary, which seemed to violate the earlier conclusion. Practically, this contradiction was attributed to the only difference in testing procedure, that is, the immediate tests administered after each viewing session in the later study did not allow students much time to internalize the knowledge. This explanation implied that "gaining benefits in terms of language acquisition appeared to be a lengthy process" (Vanderplank, 2016, p. 238–239) and, meanwhile, emphasized the importance of long-term intervention. The invalidity of on-screen texts on vocabulary learning was subsequently confirmed as a result of the low L2 proficiency in a year-long study (i.e., Muñoz et al., 2021).

Hao et al. (2021) were the representatives in the marginal third camp who believed that subtitles were more beneficial than captions in L2 vocabulary learning. They conducted another experiment that brought in dual subtitles and found that subtitles outperformed captions among the advanced learners after a 2-week intervention. However, it should be noticed that the control group without any on-screen text was the best of all, followed by the dual subtitle, the subtitle, and the caption groups in descending order, which denied the effects of on-screen texts on the whole among these advanced learners. This exceptional outcome was then ascribed to the redundancy effect (see Hao et al., 2021).

Despite the controversies in the experimental results, a tendency can still be spotted: more researchers tend to believe that subtitles show no advantage over captions in L2 vocabulary learning. A widely accepted explanation is that captions can help learners segment the speech stream in the L2 soundtrack and facilitate form-meaning mapping (Winke et al., 2010; Peters et al., 2016; Peters, 2019; Hsieh, 2020; Montero Perez, 2022), while subtitles cannot. As regards the second camp, researchers may attribute the inefficiency to learners' low proficiency or different testing procedure. However, the participants' L2 proficiency in these experiments was from beginning to intermediate. While some scholars may hold that subtitles are more suitable for beginners and captions for higher-level learners (e.g., Vanderplank, 2010), it is noted that appropriateness is the basic law in selecting learning materials and testing procedures. "Concerns about whether lower-level students can benefit from captions in the same way as upper-level learners may be more about the appropriateness of the video's complexity level for the lower-level learners rather than the appropriateness of the captioning for lower-level learners" (Winke et al., 2010, p. 80). Therefore, the difficulty of audiovisual input, as a relative concept, should be, alongside the testing procedure, adapted to match or be slightly above the learners' ability. If so, a different picture may be unfolded. To date, the role of different types of on-screen texts is yet to be further explored, especially the newly emerging dual subtitle, whether it may turn out to be the optimum type of on-screen text which incorporates



the advantages of both subtitles and captions, or a redundant type which distracts learners' limited cognitive resources.

## Effects of Different Modes of Captions

Apart from the optimum language to be displayed on the screen, the way of displaying the text is another issue of concern. Given the abovementioned advantages of captions, there has been an inclination toward an in-depth investigation into the effective way of displaying this type of on-screen text among scholars. They put forward two main approaches to make captions more accessible so as to maximize their functions: one was to reduce the captions on-screen, which resulted in a concise mode, namely, keyword caption; the other was to enrich the captions, which gave rise to the modes of glossed (full) caption, annotated caption, and enhanced caption. These modes, rising in response to their corresponding research purposes, may integrate one another in experiments and generate more variations, such as glossed keyword caption (i.e., keyword caption with corresponding L1 context-bound translation to each individual word; e.g., Hsu et al., 2013; Montero Perez et al., 2018; Teng, 2020) or enhanced and glossed caption (i.e., full caption with heightened target words and their contextual L1 translation; e.g., Hsieh, 2020). Due to the diversification of caption modes, we classify the empirical studies based on their methodology (i.e., caption modes) instead of the research results (Table 2).

The efficiency of the first approach to reduce captions was a major concern among scholars. As a result, the comparison between full captions and keyword captions has been the fundamental part, to a certain extent, in the bulk of captioning studies. Montero Perez et al. (2015) compared the effects of these two major caption modes and evaluated the (high-) intermediate proficiency learners' vocabulary increment in a comprehensive way. They found keyword captions were more robust in promoting form recognition than full captions, but helpless in terms of clip association and meaning recall. The results indicated that visual salience *via* reduction overtly directed learners' attention to the L2 word form, but failed to induce further form-meaning mapping which was regarded as one of the main benefits of full captions in L2 vocabulary learning. Treading on their heels, Teng (2019) supplemented a control group with no captions among a large number of primary school students. Overwhelming merits of full captions compared to keyword captions and no captions were observed in new words learning from all-around aspects, including form recognition, meaning recall, and meaning recognition (Teng, 2019). Though the studies yielded inverse results, they all verified the effect of full captions on constructing form-meaning connection in learners' mental lexicon.

To further enhance the accessibility of captions, Montero Perez et al. (2018) and Teng (2020) extended their research to the second approach by introducing glosses to their respective previous experimental design. Montero Perez et al. (2018) added a no caption group and a glossed keyword caption group to assess learners' vocabulary uptake. The results revealed that the glossed keyword caption group picked up most in form recognition, clip association, and meaning recall, the keyword

**TABLE 2 |** Classification of studies based on caption modes.

Caption modes	Empirical studies	Effects of caption modes (in declining order)
First direction (reduction)		
Keyword captions (KC)	Montero Perez et al., 2015 Teng, 2019	KC, FC FC, KC, 0
Second direction (addition)		
Glossed (full) captions (GFC)/ glossed keyword caption (GKC)	Montero Perez et al., 2018 Teng, 2020 Hsu et al., 2013	GKC, KC, FC, 0 GFC, FC, GKC, KC GFC, GKC, 0
Annotated captions (AC)	Aldera and Mohsen, 2013 Mohsen, 2016b	AC, FC, 0 AC, AT
Enhanced captions (EC)	Montero Perez et al., 2014b Hsieh, 2020 Cintrón-Valentín and García-Amaya, 2021 Majuddin et al., 2021	EC, KC, FC, 0 EGC, FC, EC, 0, FCNA EC, 0 EC, FC, 0

FC, Full caption; 0, No caption; AT, Annotated transcript; EGC, Enhanced and glossed caption; and FCNA, Full caption without audio.

caption and the full caption groups did not show much difference in these tests, and the no caption group was the poorest. Besides, the keyword caption group was only slightly worse than the glossed one in form recognition. The findings not only attached great importance to glosses, but also corroborated their earlier emphasis on visual salience *via* reduction (i.e., Montero Perez et al., 2015). Teng (2020), on the other hand, added one more new mode, glossed full caption, to fully tap into the competition between the full caption groups and the keyword caption groups. In line with his previous results (i.e., Teng, 2019), the full ones outperformed the keyword ones, with the scores of word form, meaning, and use ranging in a declining order: glossed full caption, full caption, glossed keyword caption, and keyword caption. It can be drawn from this study that when compared to the integrity of captions, visual salience was put in the shade in fostering vocabulary uptake, consistently contradictory to the results of Montero Perez et al. (2015, 2018). Nonetheless, the contradiction stemmed from synchronic studies. Hsu et al. (2013) observed that elementary school students in the glossed full caption and the glossed keyword caption groups shared a similar incremental pattern in their vocabulary capability in a period of 1 month, which smoothed the dispute by taking intervention duration into account.

Another line of scholars endeavored to make captions more comprehensible and conspicuous through annotations and enhancement. Aldera and Mohsen (2013) employed annotations in captioned animation to foster word learning. They assigned the high beginners to three interventional conditions: full captions, annotated captions, no captions. Annotations in this study demonstrated the utmost potential in fueling L2 word recognition and production over the short and the long term (i.e., 4 weeks). Afterward, Mohsen (2016b) compared the effects of captions and transcripts, along with annotations. The positive results of both groups confirmed the combined effects of annotations and captions/transcripts on the one hand, and the

better performance of the caption group provided additional evidence to consolidate the role of captions in L2 vocabulary development on the other hand.

Montero Perez et al. (2014b) introduced enhanced captions into their research to find the best facilitating mode. They reported that full captions, keyword captions and enhanced captions all improved word knowledge in form recognition and clip association significantly compared to the control group with no captions, thus stressed the importance of captions. Meanwhile, the results also exhibited greater influence of keyword captions and enhanced captions than full captions on meaning recognition, which aligned itself with the later findings (i.e., Montero Perez et al., 2015, 2018) that salience was superior to integrity in presenting captions. However, a more recent study (i.e., Majuddin et al., 2021), which tried to minimize the negative impact of a long lapse (4 weeks) between exposure and testing, as indicated in Vulchanova et al. (2015), by shortening the interval to 2 weeks, yielded partially different results. In their study, the college students in the full caption group and the enhanced caption group with bolded or underlined target items both outscored the control group with no captions in the form recall test immediately after the viewing. However, no advantage of salience over integrity was observed since the enhanced one did not differ much from the full one. Furthermore, the positive impact of enhancement disappeared among participants with high pretest scores in the delayed test, implying a general decline of effects of enhanced captions over a certain period. And this was also the case in Cintrón-Valentín and García-Amaya (2021) which demonstrated an obvious positive impact of enhanced captions in word recognition and production in the immediate posttest, but declined sharply in 2 weeks. Additionally, there was an even more mixed mode—enhanced and glossed caption—in audiovisual input. Hsieh (2020) conducted an experiment on low-intermediate EFL learners assigned to five groups, including no captions, full captions, full captions without audio, enhanced captions, enhanced and glossed captions, to examine the effects of each mode on L2 vocabulary improvement. The author found that the enhanced and glossed caption group surpassed the other groups in that it was not only prominent in form recognition, but also outstanding in meaning recognition and meaning recall. The one in the second place was the full caption group who did fairly well in the three tests, whereas the enhanced caption group only scored significantly higher than the other two groups in form recognition. These results indicated that enhancement and glosses could corporately contribute to learners' high involvement and form-meaning link construction and that captions could be taken as a scaffold only when synchronous audio input was available.

In sum, these experiments were conducted on the assumption that captions were beneficial to L2 vocabulary learning, and all lived up to the expectations. They provide research-based evidence to support the two approaches of releasing the potential of captions to the maximum. Specifically, though there is a debate on the effectiveness of reduction (i.e., keyword caption) compared to integrity (i.e., full caption), the former always demonstrates positive impact on vocabulary learning gains.

As for the accessibility of captions promoted by addition, glosses and annotations are proved to be facilitative, and the effectiveness of enhancement is obvious immediately after the viewing but declines as time goes by.

## Other Influencing Factors Concurring With On-Screen Texts

In reviewing the empirical literature, we noticed that some other factors did exert enormous influences on L2 vocabulary learning from audiovisual input, albeit the general positive impact of on-screen texts. These factors mainly include such learner-related ones as vocabulary size, L2 proficiency, and learning style, as well as experiment-related ones as frequency of occurrence, number of viewing, instructional strategy, and test time. They were sometimes specifically examined, and sometimes discussed as by-products in their experimental results. Additionally, due to the diversity of research in this area, more factors concerning learner, input, and test are only occasionally explored or even underexplored. **Table 3** lists these influencing factors that possibly concur with on-screen texts.

Among the learner-related factors, learners' prior vocabulary knowledge and L2 proficiency are deemed to be the key elements in deciding L2 vocabulary gains from written input (e.g., Alavi and Kaivanpanah, 2009; Lee et al., 2020). Inspired by the enlightening results in previous reading studies, scholars attempted to figure it out whether it was also the case from audiovisual input. The literature in this viewing area<sup>4</sup> has almost reached consensus on the positive role of vocabulary size: the more words a learner knows, the more gains the learner will harvest (e.g., Fievez et al., 2021). Vocabulary size was reported to greatly improve form recognition, clip association, and form recall and have even greater effect sizes than caption modes (Montero Perez et al., 2014b) and on-screen text types (Peters et al., 2016) for meaning recognition and recall. Additionally, some studies though only made a passing mention on vocabulary size, taking it as a by-product, all agreed upon its positive

<sup>4</sup>Since on-screen text is the focus of this review, some empirical studies that investigated factors influencing L2 vocabulary learning by employing audiovisual input without any on-screen text will not be considered.

**TABLE 3 |** Explored, occasionally explored, and underexplored influencing factors.

	Learner-related factors	Experiment-related factors	
		Input-related ones	Test-related ones
Explored	Vocabulary size L2 proficiency	Frequency of occurrence Number of viewing Instructional strategy	Test time
Occasionally explored	Learning style	Language distance Captioning order Content familiarity	Test modality Retention interval
Underexplored	Working memory	Duration of the viewing Number of sessions Types of video	Aspect of word knowledge Test instrument Task type

role (e.g., Montero Perez et al., 2015, 2018; Majuddin et al., 2021). The situation in L2 proficiency was a bit more complex. While proficiency was found superior to on-screen text types in predicting learners' success in both word form and word meaning uptake (Muñoz et al., 2021), there was another view that vocabulary gains bore little relation to learners' prior L2 proficiency (Frumuselu et al., 2015). Compared to the above two main factors, learning style is a marginal one in discussion. It consists of four dimensions, among which the visual/verbal learning style dimension is most suitable for audiovisual input. We are aware of only one study (i.e., Hsu et al., 2013) that tentatively took this dimension into consideration but did not discuss it in detail.

As to experiment-related factors, frequency of occurrence is one of the most popular topics inserted in the discussion of on-screen texts. Peters et al. (2016) observed an almost synchronous increase between the odds of learning a word and its frequency of occurrence, although frequency was dependent on vocabulary size. Besides, a number of studies which definitely examined the influence of frequency of occurrence confirmed its positive relation with gains in word form and meaning (e.g., Teng, 2019; Fievez et al., 2021; Muñoz et al., 2021). A similar factor which also underlines the role of repetition in vocabulary learning is the number of viewing, which was specifically examined in a study of multiword expressions learning initiated by Majuddin et al. (2021). The results, undoubtedly, showed the beneficial role of more encounters. In addition to the factors related to learning materials, two instructional strategies were employed to promote the learning outcome. One was test announcement strategy that announced an upcoming vocabulary test before the viewing so as to arouse learners' attention to the unknown words in captions, which, in turn, might enhance vocabulary gains (Montero Perez et al., 2015, 2018); and the other was advance-organizer strategy, commonly used in in-class instruction, that helped learners form a conceptual framework before viewing by providing relevant background information as well as activities (Teng, 2020). While the former strategy then turned out to be less facilitative (Montero Perez et al., 2015, 2018), the latter demonstrated great potential in helping vocabulary learning (Teng, 2020). Another factor that cannot be neglected in predicting the learning outcome is something concerning test time. Practically, this factor can be subdivided into two minor ones: the interval between the intervention and the posttest (i.e., immediate or delayed) and testing procedure of the posttest (i.e., at the end of the whole study or after each session) specifically in longitudinal studies. Some studies administered both immediate posttests and delayed tests in order to reveal the durability of the observed gains (e.g., Aldera and Mohsen, 2013; Mohsen, 2016b; Cintrón-Valentín and García-Amaya, 2021; Majuddin et al., 2021). Generally, the scores in these delayed tests were confronted with a sharp decline compared to the immediate posttests, but still higher than those without any intervention, which indicated that the vocabulary knowledge facilitated by the on-screen texts needs to be consolidated after viewing, probably through more regular encounters, to sustain its retention. Others paid attention to the long-term effect by

administering delayed tests only (e.g., Yuksel and Tanriverdi, 2009; Vulchanova et al., 2015). Though the between-group differences were not significant, the on-screen texts still showed positive impact on word learning, which provided evidence for their supportive role in L2 vocabulary retention. Empirical results also varied with the testing procedure. The typical examples were the different results derived from two 7-week longitudinal studies which were identical in every aspect except for the testing procedure, one with a tests-at-the-end format (i.e., Frumuselu et al., 2015) and the other with a tests-after-each-session format (i.e., Frumeselu, 2019), as mentioned above (see *Effects of Different Types of On-screen Texts*). The two contradictory results revealed that vocabulary learning was a process of accumulation in which learners needed time to internalize the new information they received. More longitudinal studies (e.g., Hsu et al., 2013; Muñoz et al., 2021) tended to choose the regular testing pattern after each session which could not only record the whole learning process regularly but also reflect the learning outcome more comprehensively without missing any important turning point, especially in a long-lasting experiment.

We may find that not all the factors mentioned were definitely examined in such viewing studies, and some of them were only mentioned as by-products in their discussions. Some other factors, such as test modality, language distance between the native and the target language, captioning order, content familiarity, and retention interval, were only occasionally discussed in the context of audiovisual input (e.g., Sydorenko, 2010; Winke et al., 2010, 2013; Fievez et al., 2021). And there are still a number of factors underexplored, including such input-related ones as duration of the viewing, number of sessions, types of the video (e.g., cartoons, TV series, or documentaries), and some test-related ones like aspect of word knowledge (e.g., form/meaning/use, productive/receptive, or as a whole), test instrument (e.g., Vocabulary Knowledge Scales or Vocabulary Levels Test), task type (e.g., multiple choice, cloze, or question and answer), and so forth. Besides, there is a notable lack of viewing research on learner-related factors, for example, learning style and working memory. Each of the above factors may contribute to the growth of L2 vocabulary from videos with on-screen texts, the extent to which needs to be identified through more well-controlled examinations.

## CONCLUSION AND FUTURE DIRECTIONS

Audiovisual input has greatly contributed to the development of CALL and the transformation of SLA from a book-based to a video-based activity. This review sets out to depict a comprehensive picture of the major findings from two lines of research in the past decade, with types of on-screen texts and modes of captions as their foci, to unveil the effects of videos with on-screen texts on L2 vocabulary learning. The major theories prevailing in CALL or SLA research under the audiovisual condition are introduced to improve the understanding of the designs and results of the studies.

By scrutinizing the most relevant empirical studies, this review generalizes three major findings to reflect the status quo. First, as for the helpful type (or language) of the on-screen text, captions (L2) turn out to be more robust in facilitating L2 vocabulary learning than subtitles (L1). Although a greater number of studies claimed that types of on-screen texts had little to do with the learning outcome, some defects in these studies cannot be ignored, such as the appropriateness of the difficulty of audiovisual input and the feasibility of the test procedure. What is more, there is a paucity of data about the newly emerging dual subtitles. So it is too early to draw a conclusion on the optimum type of on-screen text. Second, concerning the various modes of captions, it is more difficult to announce which one is superior to the others in aiding L2 vocabulary development. Nevertheless, some indications may be drawn from the present studies: (1) the dispute over the superiority of captioning integrity or visual salience *via* reduction has not been solved; (2) glosses and annotations which provide opportune and accessible meanings of the target words are conducive to the construction of form-meaning connection; (3) enhancement which has yielded mixed results—effective in the short term but ineffective in the long term—may be accompanied with other techniques, such as glosses and annotations to increase the overall effectiveness. Third, apart from the foci of this review (i.e., types of on-screen texts and modes of captions), some learner-related factors (e.g., vocabulary size and L2 proficiency) share similar influence on L2 word learning in audiovisual input with those in written input. And due to the nature of audiovisual input, there are quite a number of experiment-related factors to be explored, among which some factors related to learning materials, instructional strategies, and test time have already been taken into consideration, while more are underexplored and call for more specific and accurate experimental designs.

Since the initial shift of subtitles and captions to the SLA and the CALL domains, researchers have devoted themselves to investigating the functions of such on-screen texts in almost every aspect of L2 learning, especially vocabulary development. They first studied subtitles and captions separately and reported their benefits in enhancing vocabulary learning outcomes, which provided insights for the following in-depth research. Consequently, the past 10 years witnessed a tendency that a growing number of studies have endeavored to find the most beneficial type of on-screen text, namely, the most suitable language to be displayed on the screen. When the role of captions was gradually confirmed, researchers have turned to explore the rewarding way of displaying these captions (i.e., modes of captions).

This tendency implies that on-screen texts are of great potential in facilitating L2 vocabulary learning on the whole. The problem is how to fully tap into their potentials (i.e., in proper language, with suitable displaying mode, and to various groups of learners). Hence, researchers are encouraged to look more deeply into this field in the future and suggestions are provided in the following respects. First of all, as for the proper language, even though subtitles and captions have been extensively examined, dual subtitles, the newly emerging type of on-screen text which

may combine the boon of both subtitles and captions, need to be further explored. Secondly, in terms of the suitable displaying mode of captions, the superiority of integrity or visual salience *via* reduction is still open to debate. We will wait and see more studies to be engaged and expect clear-cut recommendations for L2 teaching and learning. Thirdly, regarding the various groups of learners, the bulk of the current studies were conducted in an EFL context which invariably took English as their target language, only with a few exceptions (e.g., Sydorenko, 2010; Bisson et al., 2014; Montero Perez et al., 2014b, 2015, 2018; Cintrón-Valentín and Garcíá-Amaya, 2021). The situation may be counterbalanced by introducing various second languages to English speakers (e.g., Winke et al., 2010, 2013). Moreover, the majority of the native languages belonged to the same language family—the Indo-European family, for example, Dutch, French, Spanish, and Norwegian. However, the distance between native and target languages may also influence the learning outcomes. Therefore, languages in different families (e.g., Chinese, Arabic, Finnish, and Japanese) also deserve attention. Fourthly, as to the design of the research, some high-tech devices, such as eye-tracking technique, may supplement the off-line statistics derived from the traditional test format with online performance of the learners' vocabulary process. Besides, most studies adopted a one-off intervention to tap into the role of audiovisual input in L2 vocabulary building, but it is also urgent to evaluate its long-term effect in that vocabulary learning “is an incremental process in which words should be encountered and retrieved repeatedly before they can be firmly entrenched in the mental lexicon” (Peters et al., 2016, p. 146). Fifthly, it is suggested that more studies be conducted to explore the effects of on-screen texts on the learning of formulaic sequences which comprise idioms, collocations, and other multiword units, as they are pervasive in authentic input and contribute to idiomatic L2 competence (Gholami, 2021a,b, 2022). Finally, new technologies dealing with different audiovisual input and their effectiveness need to be examined and reported timely so as to contribute to this promising area (e.g., Lin, 2021; Wu et al., 2021). Teaching with audiovisual input may turn out to be the next revolution in L2 vocabulary learning, since the ever-developing multimedia technology offers easy accessibility and simplified manipulation to teachers and learners to meet their requirements whenever and wherever possible.

## AUTHOR CONTRIBUTIONS

RW contributed to the conceptualization, investigation, and writing—original draft. LF contributed to the conceptualization and writing—review and editing, and supervision. Both authors contributed to the article and approved the submitted version.

## FUNDING

This study was funded by Ma'anshan University under Project for Cultivating Outstanding Talents in Universities.



## REFERENCES

- Alavi, S. M., and Kaivanpanah, S. (2009). Examining the role of individual differences in lexical inferencing. *J. Appl. Sci.* 9, 2829–2834. doi: 10.3923/jas.2009.2829.2834
- Aldera, A., and Mohsen, M. (2013). Annotations in captioned animation: effects on vocabulary learning and listening skills. *Comput. Educ.* 68, 60–75. doi: 10.1016/j.compedu.2013.04.018
- Alobaid, A. (2021). ICT multimedia learning affordances: role and impact on ESL learners' writing accuracy development. *Heliyon* 7:e07517. doi: 10.1016/j.heliyon.2021.e07517
- Baranowska, K. (2020). Learning most with least effort: subtitles and cognitive load. *ELT J.* 74, 105–115. doi: 10.1093/elt/ccz060
- Barón, J., and Celaya, M. L. (2022). 'May I do something for you?' The effects of audio-visual material (captioned and non-captioned) on EFL pragmatic learning. *Lang. Teach. Res.* 26, 238–255. doi: 10.1177/13621688211067000
- Birulés-Muntané, J., and Soto-Faraco, S. (2016). Watching subtitled films can help learning foreign languages. *PLoS One* 11, 1–10. doi: 10.1371/journal.pone.0158409
- Bisson, M.-J., van Heuven, W. J. B., Conklin, K., and Tunney, R. J. (2014). Processing of native and foreign language subtitles in films: an eye tracking study. *Appl. Psycholinguist.* 35, 399–418. doi: 10.1017/S0142176412000434
- Cintrón-Valentín, M. C., and García-Amaya, L. (2021). Investigating textual enhancement and captions in L2 grammar and vocabulary: An experimental study. *Stud. Second. Lang. Acquis.* 43, 1068–1093. doi: 10.1017/S0272263120000492
- Danan, M. (2004). Captioning and subtitling: undervalued language learning strategies. *Meta* 49, 67–77. doi: 10.7202/009021ar
- Fievez, I., Montero Perez, M., Cornillie, F., and Desmet, P. (2021). Promoting incidental vocabulary learning through watching a French Netflix series with glossed captions. *Comput. Assist. Lang. Learn.*, 1–26. doi: 10.1080/09588221.2021.1899244
- Frumeselu, A. D. (2019). "A friend in need is a film indeed": teaching colloquial expressions with subtitled television series" in *Using Film and Media in the Language Classroom: Reflections on Research-Led Teaching*. eds. C. Herrero and I. Vanderschelden (Bristol, Blue Ridge Summit: Multilingual Matters), 92–107.
- Frumuselu, A. D., De Maeyer, S., Donche, V., and Colon Plana, M. (2015). Television series inside the EFL classroom: bridging the gap between teaching and learning informal language through subtitles. *Linguistics Educ.* 32, 107–117. doi: 10.1016/j.linged.2015.10.001
- Gholami, L. (2021a). Incidental reactive focus on form in language classes: learners' formulaic versus nonformulaic errors, their treatment, and effectiveness in communicative interactions. *Foreign Lang. Ann.* 54, 897–922. doi: 10.1111/flan.12546
- Gholami, L. (2021b). Oral corrective feedback and learner uptake in L2 classroom: non-formulaic vs. formulaic errors. *Lang. Teach. Res.* 560. doi: 10.1177/13621688211021560
- Gholami, L. (2022). Incidental focus-on form characteristics: predicting learner uptake. Formulaic vs. non-formulaic forms. *Vigo Int. J. Appl. Linguist.* 27, 67–102. doi: 10.1017/S0272263105050163
- Hao, T., Sheng, H., Ardasheva, Y., and Wang, Z. (2021). Effects of dual subtitles on Chinese students' English listening comprehension and vocabulary learning. *Asia-Pac. Educ. Res.* doi: 10.1007/s40299-021-00601-w
- Hsieh, Y. (2020). Effects of video captioning on EFL vocabulary learning and listening comprehension. *Comput. Assist. Lang. Learn.* 33, 567–589. doi: 10.1080/09588221.2019.1577898
- Hsu, C.-K., Hwang, G.-J., Chang, Y.-T., and Chang, C.-K. (2013). Effects of video captioning modes on English listening comprehension and vocabulary acquisition using handheld devices. *J. Educ. Technol. Soc.* 16, 403–414.
- Koolstra, C. M., and Beentjes, J. W. J. (1999). Children's vocabulary acquisition in a foreign language through watching subtitled television programs at home. *Educ. Technol. Res. Dev.* 47, 51–60. doi: 10.1007/BF02299476
- Koolstra, C. M., Peeters, A. L., and Spinhof, H. (2002). The pros and cons of dubbing and subtitling. *Eur. J. Commun.* 17, 325–354. doi: 10.1177/0267323102017003694
- Lee, P.-J., Liu, Y.-T., and Tseng, W.-T. (2021). One size fits all? In search of the desirable caption display for second language learners with different caption reliance in listening comprehension. *Lang. Teach. Res.* 25, 400–430. doi: 10.1177/1362168819856451
- Lee, M., and Révész, A. (2018). Promoting grammatical development through textually enhanced captions: An eye-tracking study. *Modern Lang. J.* 102, 557–577. doi: 10.1111/modl.12503
- Lee, M., and Révész, A. (2020). Promoting grammatical development through captions and textual enhancement in multimodal input-based tasks. *Stud. Second. Lang. Acquis.* 42, 625–651. doi: 10.1017/S0272263120000108
- Lee, H., Warschauer, M., and Lee, J. H. (2020). Toward the establishment of a data-driven learning model: role of learner factors in corpus-based second language vocabulary learning. *Modern Lang. J.* 104, 345–362. doi: 10.1111/modl.12634
- Liao, S., Kruger, J. L., and Doherty, S. (2020). The impact of monolingual and bilingual subtitles on visual attention, cognitive load, and comprehension. *J. Spec. Transl.* 33, 70–98.
- Lin, P. (2021). Developing an intelligent tool for computer-assisted formulaic language learning from YouTube videos. *Recall* 34, 185–200. doi: 10.1017/S0958344021000252
- Low, R., and Sweller, J. (2014). "The modality principle in multimedia learning" in *The Cambridge Handbook of Multimedia Learning*. ed. R. E. Mayer. 2nd Edn. (New York: Cambridge University Press), 227–246.
- Lwo, L., and Lin, M. C.-T. (2012). The effects of captions in teenagers' multimedia L2 learning. *ReCALL* 24, 188–208. doi: 10.1017/S0958344012000067
- Majuddin, E., Siyanova-Chanturia, A., and Boers, F. (2021). Incidental acquisition of multiword expressions through audiovisual materials: The role of repetition and typographic enhancement. *Stud. Second. Lang. Acquis.* 43, 985–1008. doi: 10.1017/S0272263121000036
- Matielo, R., D'Ely, R. C. S. F., and Baretta, L. (2015). The effects of interlingual and intralingual subtitles on second language learning/acquisition: a state-of-the-art review. *Trab. Linguist. Apl.* 54, 161–182. doi: 10.1590/0103-18134456147091
- Mayer, R. E. (2014). "Introduction to multimedia learning" in *The Cambridge Handbook of Multimedia Learning*. 2nd Edn. (New York: Cambridge University Press), 1–24.
- Mitterer, H., and McQueen, J. M. (2009). Foreign subtitles help but native-language subtitles harm foreign speech perception. *PLoS One* 4:e7785. doi: 10.1371/journal.pone.0007785
- Mohsen, M. (2016a). The use of help options in multimedia listening environments to aid language learning: A review. *Br. J. Educ. Technol.* 47, 1232–1242. doi: 10.1111/bjet.12305
- Mohsen, M. (2016b). Effects of help options in a multimedia listening environment on L2 vocabulary acquisition. *Comput. Assist. Lang. Learn.* 29, 1220–1237. doi: 10.1080/09588221.2016.1210645
- Mohsen, M., and Mahdi, H. (2021). Partial versus full captioning mode to improve L2 vocabulary acquisition in a mobile-assisted language learning setting: words pronunciation domain. *J. Comput. High. Educ.* 33, 524–543. doi: 10.1007/s12528-021-09276-0
- Montero Perez, M. (2022). Second or foreign language learning through watching audio-visual input and the role of on-screen text. *Lang. Teach.* 55, 163–192. doi: 10.1017/S0261444821000501
- Montero Perez, M., Peters, E., Clarebout, G., and Desmet, P. (2014b). Effects of captioning on video comprehension and incidental vocabulary learning. *Lang. Learn. Technol.* 18, 118–141. doi: 10.125/44357
- Montero Perez, M., Peters, E., and Desmet, P. (2014a). Is less more? Effectiveness and perceived usefulness of keyword and full captioned video for L2 listening comprehension. *ReCALL* 26, 21–43. doi: 10.1017/S0958344013000256
- Montero Perez, M., Peters, E., and Desmet, P. (2015). Enhancing vocabulary learning through captioned video: An eye-tracking study. *Modern Lang. J.* 99, 308–328. doi: 10.1111/modl.12215
- Montero Perez, M., Peters, E., and Desmet, P. (2018). Vocabulary learning through viewing video: The effect of two enhancement techniques. *Comput. Assist. Lang. Learn.* 31, 1–26. doi: 10.1080/09588221.2017.1375960
- Montero Perez, M., Van Den Noortgate, W., and Desmet, P. (2013). Captioned video for L2 listening and vocabulary learning: a meta-analysis. *System* 41, 720–739. doi: 10.1016/j.system.2013.07.013



- Muñoz, C., Pujadas, G., and Pattemore, A. (2021). Audio-visual input for learning L2 vocabulary and grammatical constructions. *Second Lang. Res.*:26765832110157. doi: 10.1177/02676583211015797
- Neuman, S. B., and Koskinen, P. (1992). Captioned television as comprehensible input: effect of incidental word learning from context for language minority students. *Read. Res. Q.* 27, 94–106. doi: 10.2307/747835
- Paivio, A. (1969). Mental imagery in associative learning and memory. *Psychol. Rev.* 76, 241–263. doi: 10.1037/h0027272
- Paivio, A. (1990). *Mental Representations: A Dual Coding Approach*. New York: Oxford University Press.
- Paivio, A. (2010). Dual coding theory and the mental lexicon. *Ment. Lex.* 5, 205–230. doi: 10.1075/ml.5.2.04pai
- Pattemore, A., and Muñoz, C. (2020). Learning L2 constructions from captioned audio-visual exposure: The effect of learner-related factors. *System* 93:102303. doi: 10.1016/j.system.2020.102303
- Peters, E. (2019). The effect of imagery and on-screen text on foreign language vocabulary learning from audio-visual input. *TESOL Q.* 53, 1008–1032. doi: 10.1002/tesq.531
- Peters, E., Heynen, E., and Puimège, E. (2016). Learning vocabulary through audio-visual input: The differential effect of L1 subtitles and captions. *System* 63, 134–148. doi: 10.1016/j.system.2016.10.002
- Pujadas, G., and Muñoz, C. (2020). Examining adolescent EFL learners' TV viewing comprehension through captions and subtitles. *Stud. Second. Lang. Acquis.* 42, 551–575. doi: 10.1017/S0272263120000042
- Sadoski, M., and Paivio, A. (2000). *Imagery and Text: A Dual Coding Theory of Reading and Writing*. Mahwah, NJ: Lawrence Erlbaum Associates, Inc.
- Sweller, J., van Merriënboer, J. J. G., and Paas, F. (2019). Cognitive architecture and instructional design: 20 years later. *Educ. Psychol. Rev.* 31, 261–292. doi: 10.1007/s10648-019-09465-5
- Sydorenko, T. (2010). Modality of input and vocabulary acquisition. *Lang. Learn. Technol.* 14, 50–73. doi: 10125/44214
- Teng, F. (2019). Incidental vocabulary learning for primary school students: The effects of L2 caption type and word exposure frequency. *Aust. Educ. Res.* 46, 113–136. doi: 10.1007/s13384-018-0279-6
- Teng, M. F. (2020). Vocabulary learning through videos: captions, advance-organizer strategy, and their combination. *Comput. Assist. Lang. Learn.* 35, 518–550. doi: 10.1080/09588221.2020.1720253
- Teng, M. F. (2021). *Language Learning Through Captioned Videos: Incidental Vocabulary Acquisition*. New York: Taylor & Francis Group.
- Vanderplank, R. (2010). Déjà vu? A decade of research on language laboratories, television and video in language learning. *Lang. Teach.* 43, 1–37. doi: 10.1017/S0261444809990267
- Vanderplank, R. (2016). 'Effects of' and 'effects with' captions: how exactly does watching a TV programme with same-language subtitles make a difference to language learners? *Lang. Teach.* 49, 235–250. doi: 10.1017/S0261444813000207
- Vulchanova, M., Aurstad, L. M. G., Kvitnes, I. E. N., and Eshuis, H. (2015). As naturalistic as it gets: subtitles in the English classroom in Norway. *Front. Psychol.* 5:1510. doi: 10.3389/fpsyg.2014.01510
- Wang, A., and Pellicer-Sánchez, A. (2022). Using bilingual subtitles while watching videos to facilitate vocabulary learning. *Lang. Learn.* 7:12495. doi: 10.1111/lang.12495
- Wilkins, D. A. (1972). *Linguistics in Language Teaching*. London: Edward Arnold.
- Winke, P., Gass, S., and Sydorenko, T. (2010). The effects of captioning videos used for foreign language listening activities. *Lang. Learn. Technol.* 14, 65–86. doi: 10125/44203
- Winke, P., Gass, S., and Sydorenko, T. (2013). Factors influencing the use of captions by foreign language learners: An eye-tracking study. *Modern Lang. J.* 97, 254–275. doi: 10.1111/j.1540-4781.2013.01432.x
- Wisniewska, N., and Mora, J. C. (2020). Can captioned video benefit second language pronunciation? *Stud. Second. Lang. Acquis.* 42, 599–624. doi: 10.1017/S0272263120000029
- Wong, S. W. L., Lin, C. C. Y., Wong, I. S. Y., and Cheung, A. (2020). The differential effects of subtitles on the comprehension of native English connected speech varying in types and word familiarity. *SAGE Open* 10, 215824402092437–215824402092413. doi: 10.1177/2158244020924378
- Wu, W.-C. V., Lin, I.-T. D., Marek, M. W., and Yang, F.-C. O. (2021). Analysis of English idiomatic learning behaviors of an audio-visual mobile application. *SAGE Open* 11:215824402110168. doi: 10.1177/21582440211016899
- Yeldham, M. (2018). Viewing L2 captioned videos: What's in it for the listener? *Comput. Assist. Lang. Learn.* 31, 367–389. doi: 10.1080/09588221.2017.1406956
- Yuksel, D., and Tanriverdi, B. (2009). Effects of watching captioned movie clip on vocabulary development of EFL learners. *Turkish Online J. Educ. Technol.* 25, 525–532. doi: 10.1016/j.tate.2008.09.012

**Conflict of Interest:** The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

**Publisher's Note:** All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Copyright © 2022 Wei and Fan. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.



# Teaching Academic Words With Digital Flashcards: Investigating the Effectiveness of Mobile-Assisted Vocabulary Learning for University Students

Ismail Xodabande<sup>1\*</sup>, Yasaman Iravi<sup>2</sup>, Behzad Mansouri<sup>3</sup> and Hoda Matinparsa<sup>4</sup>

<sup>1</sup>Department of Foreign Languages, Kharazmi University, Tehran, Iran, <sup>2</sup>Department of Foreign Languages, Imam Khomeini International University, Qazvin, Iran, <sup>3</sup>Research and Innovation Department, Lakeshore Foundation, Birmingham, AL, United States, <sup>4</sup>Department of Foreign Languages, Islamic Azad University, Qazvin, Iran

## OPEN ACCESS

### Edited by:

Ehsan Rassaei,  
Majan University College, Oman

### Reviewed by:

Ali Soyoof,  
Monash University, Australia  
Parisa Abdolrezapour,  
Salman Farsi University of Kazerun,  
Iran

### \*Correspondence:

Ismail Xodabande  
ismail.kh.tefl@gmail.com

### Specialty section:

This article was submitted to  
Language Sciences,  
a section of the journal  
Frontiers in Psychology

**Received:** 10 March 2022

**Accepted:** 09 May 2022

**Published:** 14 June 2022

### Citation:

Xodabande I, Iravi Y, Mansouri B and  
Matinparsa H (2022) Teaching  
Academic Words With Digital  
Flashcards: Investigating the  
Effectiveness of Mobile-Assisted  
Vocabulary Learning for  
University Students.  
Front. Psychol. 13:893821.  
doi: 10.3389/fpsyg.2022.893821

The current study explored the effects of using digital flashcards (DFs) and mobile devices on learning academic vocabulary. The participants were 86 university students majoring in Psychology in two experimental conditions and one control group. A list of 361 core academic words frequently used in Psychology was taught to the participants using different materials, and the learning outcomes were compared across the three groups. Accordingly, the participants in the experimental group 1 ( $N = 31$ ) used a DF application (i.e., NAWL builder), participants in the experimental group 2 ( $N = 30$ ) used traditional materials (i.e., paper flashcards), and those in the control group were given a list of target words with their definitions. Receptive knowledge of the target words was tested before and after the treatment, and the learning outcomes were compared across the groups using one-way between-groups ANOVA. The findings of the study indicated that using DFs enhanced students' engagement with learning their discipline-specific academic vocabulary and that experimental group 1 outperformed those participants in other learning conditions. The findings add to the existing literature on mobile-assisted vocabulary learning and provide empirical support for the effectiveness of such platforms for learning academic vocabulary. The implications of the study were discussed in terms of the affordances provided by DFs on mobile devices and corpus-based word lists for informing vocabulary learning components in teaching English for Academic Purposes (EAP).

**Keywords:** mobile-assisted vocabulary learning, academic vocabulary, digital flashcards, corpus-based language teaching, EAP

## INTRODUCTION

There is a consensus that learning a language is highly contingent upon mastery of its vocabulary (Nation, 2013; Webb and Nation, 2017). Vocabulary knowledge has been viewed as “the key type of knowledge necessary for any language use, because if words to express concepts are not known, all syntactic and discourse knowledge is of little use” (Schmitt et al., 2021, p. 10). The importance given to the development of vocabulary knowledge as an indicator of success

or failure in any form of second or foreign language (L2) learning has contributed to pedagogical advancement specifically focused on enhancing the quality and quantity of words learned in an L2. For example, the development of vocabulary word lists such as General Service List (GSL; West, 1953), Academic Word List (Coxhead, 2000), and Knowledge-based Vocabulary Lists (Schmitt et al., 2021) all has been confirming the fact that knowing the words of a language and learning how to use them are probably the most essential aspect of the literacy in language learning in general and L2 learning in particular (Yang and Coxhead, 2020). Accordingly, finding effective strategies to facilitate the development of lexical competence among L2 learners remained a worthwhile research agenda.

Research pertinent to English vocabulary instruction has also shown that learning English words and developing L2 vocabulary knowledge either incidentally or intentionally could be a daunting task for learners situated in English as a foreign language (EFL) contexts (Laufer, 1996; Webb and Chang, 2012; Honzard and Soyooof, 2020). The whole process could also be accompanied by excessive pressure and extra load when it comes to learning words and vocabulary items not commonly found in everyday conversations (Coxhead, 2018, 2019; Yüksel et al., 2020). In other words, learning words known as technical and academic vocabulary could seem too overwhelming to learners that they may not actively participate in the process of learning. To ease some of the pressure exerted on both L2 teachers and learners, scholars have called for the inclusion and expansion of new technologies and findings of corpus linguistics research in L2 vocabulary teaching and learning (Ma, 2017; Coxhead, 2018, 2019; Ma and Mei, 2021; Soyooof et al., 2021). In this regard, the current study aimed to investigate the effectiveness of mobile devices and digital flashcard applications in learning academic words that are frequently used in the field of Psychology. The study contributes to the expanding body of knowledge on mobile-assisted vocabulary learning (Lin and Lin, 2019), and the findings might inform academic vocabulary instruction in English for Academic Purposes (EAP) programs (Soyooof et al., 2022).

## LITERATURE REVIEW

### Academic Vocabulary

Generally operationalized as the words used more frequently in academic texts and not in non-academic texts (Nation, 2013; Coxhead, 2019), academic vocabulary refers to a class of medium-frequency words (i.e., beyond high-frequency or general service vocabulary) that are used mainly for describing abstract ideas and processes in the scientific discourse and rhetorical organization of academic texts (Paquot, 2010; Coxhead, 2018). Recent corpus-based studies pointed to a considerably large coverage for academic vocabulary ranging from 6 to 14% in academic texts (Coxhead, 2000; Browne et al., 2013; Gardner and Davies, 2014), and it has been argued that knowledge of these words is crucial for understanding and producing academic writing and more generally academic literacy development (Coxhead and Byrd, 2007; Nagy and Townsend, 2012;

Gardner and Davies, 2014). Moreover, research indicated that academic vocabulary poses major challenges in terms of the learning burden not only for English as second/foreign language (ESL/EFL) learners but also for native English-speaking students (Evans and Morrison, 2010, 2011; Spencer et al., 2017). Hence, given their importance and crucial role in academic discourse, a number of core academic word lists have been developed for setting vocabulary learning goals for university students (e.g., Coxhead, 2000; Browne et al., 2013; Gardner and Davies, 2014).

The Academic Word List (AWL; Coxhead, 2000) developed more than two decades ago, has remained a major resource for vocabulary instruction, materials development, and vocabulary assessment in EAP programs (Coxhead, 2011; McLean and Kramer, 2015). Nonetheless, a growing number of studies investigating the coverage of the AWL in different academic corpora started to challenge its position as the predominant source of core academic vocabulary relevant to a wide range of disciplines (Hyland and Tse, 2007; Gardner and Davies, 2014; Masrai and Milton, 2018). More specifically, the AWL has been criticized on various grounds including (1) the use of outdated GSL for defining general service vocabulary (Gardner and Davies, 2014), (2) containing a large number of general or only marginally academic words (Masrai and Milton, 2018), (3) the variation in the coverage of the list across disciplines (Liu and Han, 2015), and (4) using word families as the unit of counting vocabulary items which limits its pedagogical value (Gardner and Davies, 2014). In recent years, two core academic word lists, namely the New Academic Word List (NAWL; Browne et al., 2013), and the Academic Vocabulary List (AVL; Gardner and Davies, 2014), have been developed in response to the limitations associated with the AWL. These new lists showed considerable improvements in terms of their coverage in academic texts and also employed more pedagogically useful units (i.e., lemma and flemma; Brown et al., 2020) in operationalizing academic vocabulary. For example, Browne et al. (2013) developed the NAWL that contains 960 words based on a carefully selected academic corpus with 288 million words as part of the Cambridge English Corpus (CEC). General service and high-frequency vocabulary items accounted for 86% of the CEC, and the NAWL increased this coverage to around 92%. In this regard, learning the vocabulary items in the NAWL is of significant importance for university students and might be regarded as a more appropriate vocabulary learning goal. Mastery of the items in these lists facilitates achieving the minimum comprehension threshold for understanding academic discourse (Schmitt et al., 2011; Laufer, 2013).

### Incidental and Deliberate Vocabulary Learning

Developments in vocabulary knowledge in first language (L1) learning generally occur incidentally, which means that there is no conscious or explicit attention to learning words themselves, but the focus is on meaning in communicative interactions. In other words, incidental vocabulary learning has been regarded

as a by-product of communication that is associated with some sort of meaningful input (Nation, 2013). A crucial factor in incidental vocabulary development then is the amount of input, as more input increases the chances of encountering new words and boosts the likelihood of picking the meaning from the context (Webb and Nation, 2017). However, in learning additional languages beyond the L1, creating the facilitative conditions for incidental vocabulary learning is not easily possible. More specifically, language learners in most EFL contexts have very limited exposure to the target language beyond the classroom which significantly impacts incidental vocabulary learning. Accordingly, there is a growing consensus that for L2 learners, deliberate and intentional learning accounts for most developments in vocabulary knowledge (Laufer, 2005). Research in this area clearly pointed to higher gains in intentional vocabulary learning compared to incidental learning conditions for L2 learners (Webb and Nation, 2017).

Intentional vocabulary learning for L2 learners might be undertaken in different ways. Besides using available resources such as dictionaries and course books in the classroom, a range of approaches including using flashcards, learning from word lists, writing tasks, serious games, and fill-in-the-blanks activities contribute significantly to vocabulary learning (Webb et al., 2020; Li and Hafner, 2022; Soyoof et al., 2022). In this regard, one of the most effective approaches for intentional learning, paper, and digital flashcards has been employed extensively to augment vocabulary learning among EFL learners (Nakata, 2019). In a recent study, Li and Hafner (2022) explored the impacts of using word cards on mobile devices to improve receptive and productive vocabulary knowledge among Chinese EFL learners. The findings indicated that although both digital and traditional flashcards contributed significantly to vocabulary knowledge development, digital flashcards produced better learning outcomes. Moreover, recent developments in digital technologies created appropriate conditions for facilitating intentional vocabulary development in the form of self-directed learning outside the classroom. For example, Xodabande et al. (2022) investigated self-directed and intentional vocabulary learning among Iranian EFL students with paper and digital flashcards, and their findings indicated that such strategies hold considerable potential to shortcut the long-term process of vocabulary learning.

## Mobile-Assisted Vocabulary Learning

Mobile-assisted language learning (MALL) has attracted considerable attention in recent years (e.g., Dashtestani, 2016; Hwang and Fu, 2019; Chen et al., 2020; Nazari and Xodabande, 2020, 2021; Rassaei, 2020, 2021; Burston and Giannakou, 2021; Dashtestani and Hojatpanah, 2021), and a large number of studies investigated the use of different delivery mediums and environments (such as SMS, MMS, and mobile applications) for learning ESL/EFL vocabulary (Mahdi, 2017; Lin and Lin, 2019). Overall, the findings from this growing body of knowledge indicated that the use of SMS/MMS and mobile applications including context-aware, gaming, and researcher-designed platforms contributed significantly to improvements in EFL learners' vocabulary knowledge (Mahdi, 2017; Xodabande, 2017;

Lin and Lin, 2019; Soyoof et al., 2021). Additionally, a growing number of studies are investigating the affordances of digital flashcards (DFs) in the form of mobile applications for vocabulary learning (e.g., Nakata, 2019; Seibert Hanson and Brown, 2020; Xodabande and Atai, 2020; Yüksel et al., 2020; Zhang et al., 2021; Xodabande et al., 2022). In this regard, although the number of studies focusing on teaching general vocabulary is increasing, research on using such platforms (e.g., Anki) for facilitating the development of academic or technical vocabulary knowledge remained limited (Honzard and Soyoof, 2020; Yüksel et al., 2020). As such applications provide opportunities for meaningful repetition of the target vocabulary items and scaffold the learning experience through a number of multimedia features, their integration into language teaching programs seems to be especially promising for teaching academic vocabulary (Mansouri and Mantero, 2019; Xodabande and Atai, 2020). Additionally, by making use of digital flashcards, language teachers can create targeted content and opportunities to "support learners' self-directed study efforts and help them consolidate [their] vocabulary knowledge" (Yüksel et al., 2020, p. 2), thus elevating learners' agency in keeping themselves accountable for the progress in learning.

Some studies investigated the use of mobile devices for learning academic and technical vocabulary among university students. In a quasi-experimental study, Yüksel et al. (2020) compared the effectiveness of DFs (i.e., Quizlet) and wordlists on 57 undergraduate pharmacy students' technical word learning over 10 weeks by collecting data from the pre-treatment survey, two technical vocabulary tests, and the post-treatment survey. The results pointed to the high potential of DFs in technical vocabulary learning. Furthermore, learners provided an optimistic perspective on using DFs. In another study, Kohnke et al. (2019) developed an application (i.e., *Excel@EnglishPolyU*) and two vocabulary-based English language learning games for learning business vocabulary. The researchers then investigated the business vocabulary retention of 51 undergraduate students at a university in Hong Kong. Analysis of data revealed the positive impacts of mobile-gamified applications in vocabulary knowledge development. Similarly, Kohnke et al. (2020) explored the vocabulary retention of 159 ESL learners from four disciplines at Hong Kong University using an in-house mobile application specially designed to build a repertoire of field-specific academic words. Their findings from the analysis of pre- and post-tests including 120 vocabulary items revealed the beneficial impacts of mobile apps on field-specific word learning and retention. Honzard and Soyoof (2020) compared the effectiveness of using mobile apps and serious games on English word retention among 90 Iranian EFL learners. Placing participants in two groups, the researchers conducted pre-test, post-test, and delayed post-test and found that both approaches were influential in enhancing participants' word retention with serious games having an edge over a mobile application. The authors argued for the inclusion of using games besides the conventional mobile applications commonly used for vocabulary learning and teaching.

With respect to learning academic vocabulary, Dizon (2016) probed the effectiveness of DF Quizlet in vocabulary development.



The participants were nine EFL students in a Japanese university who studied Coxhead's (2001) general academic vocabulary list (AWL) over 10 weeks, and the findings of pre- and post-tests indicated that students gained considerably from the DF application. Moreover, Xodabande and Atai (2020) studied the impacts of a mobile application on self-directed learning of academic vocabulary among 38 Iranian university students. The participants were divided into experimental and control groups, and the study adopted a pre-, post-, and delayed post-test design to investigate the effects of mobile-assisted vocabulary learning in the long run. The participants in the experimental group used a flashcard mobile application to learn vocabulary items from AWL (Coxhead, 2000), and those in the control group used traditional materials for learning the same vocabulary items. Although the findings of the study showed improvements for both experimental and control groups in terms of academic vocabulary knowledge, the impacts on the experimental group were significantly higher than that of the control group. The study highlighted the potential of mobile applications for learning academic vocabulary. Similar findings were also reported by Ashcroft et al. (2018), as they compared the effect of DFs and paper flashcards on general academic vocabulary development at various English proficiency levels. Despite this emerging evidence on the benefits of using digital flashcards for learning academic and technical vocabulary, there are recent calls for more thorough and long-term intervention-based studies to comprehensively examine the impacts of utilizing such technologies on L2 vocabulary learning.

## THE PRESET STUDY

Given the importance of academic vocabulary for university students and the positive learning outcomes reported for mobile-assisted vocabulary learning, the current study aimed to explore the impacts of using DFs for teaching academic vocabulary to Psychology students. The study also compared the learning outcomes attained from using DFs to paper flashcards and word lists. The study is significant since it addresses a number of gaps in the literature. First, as needs analysis studies indicated, academic vocabulary knowledge featured high among the language learning needs of Psychology students (e.g., Atai and Hejazi, 2019). In this regard, finding effective strategies to facilitate academic vocabulary learning contributes to their academic literacy and professional identity development. Second, as highlighted above, studies investigating the impacts of mobile-assisted vocabulary learning on specialized vocabulary (i.e., academic and technical) remained limited (Honzard and Soyoo, 2020; Yüksel et al., 2020), and there is a need for further empirical research to understand both short- and long-term impacts of such interventions. Third, previous research indicated there is a considerable disciplinary variation in the way items from corpus-based word lists (such as AWL) are used in academic discourse (Hyland and Tse, 2007). As a result, teaching all items in a core academic wordlist for students in a particular field of study is not practical as many words in such lists are not relevant to their vocabulary learning needs. Fourth, with

the expanding “centrality of English as a *lingua franca*” in all academic disciplines, there is a need for exploring new avenues for university students “to learn English in contexts that are meaningful to them” (Soyoo et al., 2022, p. 5). Accordingly, the present study used the results of a large corpus-based study of Psychology texts (Valizadeh and Xodabande, 2021) for selecting target academic words and connected findings from a corpus-based study of specialized texts to mobile-assisted vocabulary learning. The following research question was investigated:

Does using DFs on mobile devices result in enhanced learning outcome in teaching academic vocabulary?

## MATERIALS AND METHODS

### Participants

The participants of the current study were 86 Iranian university students (49 females, 37 males) majoring in Psychology, recruited through purposive sampling. The following criteria were implemented in selecting the participants: nationality (i.e., Iranian), education level (i.e., Psychology major), and limited or no prior experience in living in an English-speaking country. The mean age of the participants was 22, and the majority were at the intermediate level in English based on (1) responses to the self-report proficiency measure, and (2) the results of the Cambridge Placement Test (Test Your English, 2022). The test is an online instrument with 25 multiple-choice format questions, which is used as a quick placement test for English language learners. It takes around 10 to 15 min to complete the test. At the time of the study, the participants were taking the course “English for Psychology students” as part of their 4-credit English for Specific Purposes (ESP) education. The course aimed to familiarize the students with reading disciplinary texts in Psychology. The participants were randomly assigned to three learning conditions. The experimental group 1 ( $N = 31$ ) used digital flashcards on their mobile devices, the experimental group 2 ( $N = 30$ ) used paper-based flashcards, and the control group ( $N = 25$ ) was given a list of target vocabulary items with associated definitions. All participants in the experimental group 1 owned smartphone or tablet devices for installing and using vocabulary-learning applications. The study adhered to ethical considerations in educational research by obtaining informed consent from participants and ensuring the confidentiality of the collected data.

### Materials and Instruments

#### NAWL Builder Application

The study used New Academic Word List (NAWL; Browne et al., 2013) as a source for academic vocabulary in English. Accordingly, those participants in experimental group 1 installed the NAWL builder flashcards (EFL Technologies, 2017) to learn 361 academic words frequently used in Psychology (Valizadeh and Xodabande, 2021). This application is selected for this study for several reasons. First, it is freely accessible from the Google Play Store and Apple's App Store for Android and iOS platforms. Second, using a built-in spaced repetition system



(Kornell, 2009), the application facilitates the learning of vocabulary items in the NAWL list (Browne et al., 2013). Third, the NAWL builder employs a set of simple tools for learning vocabulary which makes it easy to use for language learners and keeps a detailed record of the users' progress in vocabulary learning that can be emailed to the teacher. Moreover, the application uses simple English in the definitions given for academic vocabulary and provides part of speech information and North American pronunciation for the target words. In order to compare learning gains from different materials, the experimental group 2 was given ready-made paper flashcards for learning the same words that contained word form and related part of speech information on one side of the card, and simple definitions on the other side. The content of these word cards was similar to the cards in NAWL builder, except for the audio component of words. The control group was given a list of 361 academic words with their part of speech information and definitions.

### Vocabulary Tests

In order to test the participants' vocabulary knowledge before and following the treatment, two measures of receptive knowledge of academic vocabulary were employed. In this regard, New Academic Word List Test (NAWL; Stoeckel and Bennett, 2020) was used as the first measure, which is a standard and validated diagnostic test of written receptive knowledge of vocabulary items in the NAWL. The NAWLT contains 40 items in multiple-choice format, and short sentences containing the target word in a natural but non-defining context are provided in the questions. The development of the NAWLT items was based on sound specifications, and the test in general shows high reliability (Cronbach's  $\alpha=0.75$ ). Moreover, two Vocabulary Knowledge Tests (VKT) each containing 60 multiple-choice items were developed to test the knowledge of 361 academic words frequently used in Psychology texts. For designing these tests, 120 items out of 361 words were selected randomly and assigned to two sets using the research randomizer website. The distractors in these tests were selected from simple definitions provided for NAWL items. The reliability of the developed test was acceptable (Cronbach's  $\alpha=0.83$ ), and its validity was examined in relation to the NAWLT in a pilot testing session on a similar sample ( $N=20$ ), and the test demonstrated acceptable concurrent validity (Frey, 2018) with an established instrument.

### Procedures and Data Analysis

At the start of the academic semester, the participants' vocabulary knowledge was tested using the above-mentioned measures. This initial assessment of the vocabulary learning needs was followed by a one-hour training session for all participants on vocabulary learning strategies with focusing on digital flashcards and word lists. Then, the participants in the experimental group 2 and the control group were given ready-made flashcards and the word list, respectively. The participants in experimental group 1 installed the NAWL builder application and received instructions for selecting the 361 academic words

that are relevant to Psychology. To this end, printed copies of the list of the frequently used NAWL items in Psychology (Valizadeh and Xodabande, 2021) which were sorted in alphabetical order were given to the participants, and they were asked to select the vocabulary items in the list<sup>1</sup>.

As part of their ESP course requirements, the participants in the two experimental groups were asked to spend at least 50 min every week (10–15 min per day) to study the target words (around 25 words per week) with the assigned materials over the course of an academic semester (i.e., 15 weeks). Academic vocabulary learning accounted for 30% of the overall course grade for the participants in the experimental groups and not for the control group. At the end of the academic semester, the participants' vocabulary knowledge was tested again, to compare the learning outcomes across three learning conditions. The data obtained *via* vocabulary knowledge tests were analyzed using IBM SPSS (version 25) for both descriptive and inferential statistics. One-way between-groups ANOVA was performed to compare the scores on academic vocabulary tests on pre- and post-test.

## RESULTS

**Table 1** provides a summary of descriptive statistics for the results of the two tests, namely the NAWLT and VKT obtained by the participants on pre-test. As is shown in the table, the mean values calculated for both tests are consistent across the three groups. The total mean value for the scores was 11.73 ( $SD=2.11$ ) for the NAWLT, and the performances of the three groups were largely similar. With respect to the VKS, the total mean score for the participants was 20.28 ( $SD=2.86$ ).

In order to see if the observed variation in the scores obtained by the three groups on the pre-test is statistically significant, a one-way between-groups ANOVA was conducted. **Table 2** shows the results of Levene's test for homogeneity of variances that investigates whether the observed variance in the scores is the same for the three groups. Since the test returned a non-significant results, the homogeneity assumption of variance has not been violated for the scores obtained on the pre-test.

The results of the one-way between-groups analysis of variance (**Table 3**) revealed that the observed differences in the scores for the NAWLT,  $F(2, 83)=0.67$ ,  $p=0.512$ , and the VKT,  $F(2, 83)=0.98$ ,  $p=0.38$ , were not statistically significant. Accordingly, the results indicated that prior to the treatment, the three groups were similar in terms of their receptive knowledge of the 361 target academic words.

The results of the descriptive statistics for the scores obtained on the post-test are summarized in **Table 4**. Unlike the participants' performances on NAWLT and VKT on the pre-test, the post-test results show different learning outcomes for the three learning conditions. With respect to the NAWLT, the experimental group 1 obtained higher scores ( $M=20.42$ ,

<sup>1</sup>The NAWL software can be accessed at: <https://apps.apple.com/us/app/nawl-builder-multilingual/id1097693935>.

**TABLE 1** | Descriptive statistics for pre-test results.

		<i>N</i>	Mean	Std. deviation	Std. error	95% confidence interval for mean	
						Lower bound	Upper bound
NAWLT	EXP 1	31	11.87	2.232	0.401	10.83	11.69
	EXP 2	30	11.93	2.033	0.371	11.17	12.69
	CON	25	11.32	2.076	0.415	10.46	12.18
	Total	86	11.73	2.111	0.228	11.28	12.19
VKT	EXP 1	31	20.23	2.997	0.538	19.13	21.33
	EXP 2	30	20.80	2.578	0.471	19.84	21.76
	CON	25	19.72	3.007	0.601	18.48	20.96
	Total	86	20.28	2.860	0.308	19.67	20.89

$SD=3.68$ ), followed by the experimental group 2 ( $M=16.20$ ,  $SD=3.15$ ) and the control group ( $M=13.12$ ,  $SD=3.00$ ). As for the VKT, the post-test scores indicated a similar pattern, as the experimental group 1 obtained better scores ( $M=41.00$ ,  $SD=4.75$ ). The participants in the experimental group 2 had higher scores ( $M=34.73$ ,  $SD=5.89$ ) compared to the control group ( $M=30.92$ ,  $SD=6.99$ ).

In order to proceed with analyzing the data for inferential statistics, Levene's test for homogeneity of variances was conducted prior to ANOVA, and the results (Table 5) indicated that the assumption of homogeneity of variance has not been violated for the scores obtained on post-test. Additionally, the results of the one-way between-groups analysis of variance (Table 6) revealed that the observed differences in the scores for the NAWLT,  $F(2, 83)=34.33$ ,  $p<0.001$ ; eta squared=0.45), and the VKT,  $F(2, 83)=22.63$ ,  $p<0.001$ ; eta squared=0.35, were statistically significant. Accordingly, the results indicated that after the treatment, the three groups were different in terms of their receptive knowledge of the 361 target academic words. The effect size of the observed differences for both measures was very large based on the criteria proposed by Cohen (1988).

Finally, in order to compare and contrast the scores obtained by the three groups on the post-test, a series of pairwise comparisons were conducted (Table 7). The results revealed that the experimental group 1 that used DFs for learning academic vocabulary outperformed the experimental group 2 (mean differences: NAWLT=4.22, VKT=6.27,  $p<0.001$ ) and the control group (mean differences: NAWLT=7.29, VKT=10.08,  $p<0.001$ ). Moreover, the participants in experimental group 2 that used paper flashcards for vocabulary learning outperformed the control group (mean differences: NAWLT=3.08, VKT=3.81,  $p<0.001$ ).

## DISCUSSION

The present study investigated the impacts of using DFs on mobile devices for learning academic vocabulary by university students and compared the learning outcomes to traditional materials including paper flashcards and word lists. The results indicated that mobile-assisted vocabulary learning using digital flashcards with built-in spaced repetition technology improved

participants' academic vocabulary knowledge significantly from pre-test to the post-test, and that the participants in the experimental group outperformed the other groups on both measures of academic vocabulary knowledge. These findings are congruent with earlier studies that reported positive learning outcomes for mobile-assisted vocabulary learning (Dizon, 2016; Ashcroft et al., 2018; Kohnke et al., 2019, 2020; Xodabande and Atai, 2020; Yüksel et al., 2020). A close examination of the results obtained on pre- and post-tests on VKT (Tables 1, 4) revealed that prior to the treatment, the participants of the study were familiar with about 33% of the 361 academic words (around 120 items) that are frequently used in their field of study. However, after the semester-long treatment/instruction, the participants in the experimental group 1 learned around 126 more items and their test results pointed to achieving around 68% mastery over the target items (35% improvement). Additionally, the participants in experimental group 2 learned around 56% of the target items (23% improvement), and those in the control group learned 51% of the items (18% improvement). Accordingly, although the interventions were not effective in teaching all 361 words, considerable improvements in the vocabulary knowledge of the participants in the experimental learning conditions point to the effectiveness of explicit focus on vocabulary learning in general and the relative advantage of mobile-assisted learning in the target items in particular. It has been argued that the integration of digital technologies into language education positively impacts learners' motivation (Stockwell, 2013), and the motivational dimension of mobile-assisted vocabulary learning might be considered the key factor explaining the significant learning outcomes in the experimental group 1. Additionally, studies indicated learning vocabulary items alongside multimedia features such as pictures and audio pronunciation simultaneously improve learning outcomes (Rasti-Behbahani and Shahbazi, 2020), which partly explains better learning outcomes for digital flashcards observed in this study. Another factor that further contributed to overall improvements in the test scores might be the integration of the academic vocabulary learning component into the ongoing ESP course for the participants, which resulted in increased engagement with materials and spending more time and effort for learning target words.

Moreover, the treatment in the form of mobile-assisted vocabulary learning lasted for a semester, and the findings of the study

supported the long-term effectiveness of mobile devices and DFs in teaching academic vocabulary to university students. This is also congruent with the limited findings in the literature that reported long-term positive outcomes for mobile-assisted academic vocabulary learning (Lin and Lin, 2019). As the majority of previous studies on mobile-assisted language learning was conducted in short time periods (Chwo et al., 2018; Lin and Lin, 2019), these results are significant as they add to the growing body of knowledge in mobile-assisted learning of general and specialized vocabulary. The observed long-term effectiveness of mobile-assisted vocabulary learning might be attributed to a number of factors. First, since the target words were selected from a corpus-based study of specialized texts in Psychology, the vocabulary items were highly relevant to the participant's field of study, which might have resulted in their increased motivation and learning effort. Second, as the NAWL builder app has a built-in spaced

repetition feature, the participants learned and practiced academic vocabulary systematically and efficiently. Additionally, the availability of mobile devices and learning materials for the participants facilitated anytime and anywhere learning experience (Lin and Lin, 2019) which promoted learning outcomes over time.

Moreover, given the relatively large number of items in core academic wordlists, focusing on those words that are more frequent in a given field of study brings better learning outcomes for some reasons. First, reducing the number of target vocabulary items makes the list more manageable for students to study the vocabulary items with DFs in one or two semesters. Second, as the items are highly relevant to their discipline and professional identity, university students might be more motivated to invest time and effort in learning discipline-related/specific academic words. Third, as vocabulary instruction receives insufficient attention in language classes (Webb and Nation, 2017), such a fine-tuned approach allows instructors and students to use valuable classroom and self-study time for focusing on the most important academic vocabulary items. Finally, although the findings of the study pointed to the long-term effectiveness of the DFs in academic vocabulary learning, obtaining slightly lower (but statistically significant) scores on the delayed post-test by the participants shows that without reinforcing the developing knowledge of vocabulary items, the learning gains might be lost over time. As a result, there is a need for systematic review and passive or active use of learned items through academic reading and writing.

## CONCLUSION AND IMPLICATIONS

The current study pursued two main goals: (a) exploring the effect of using DFs on learning academic vocabulary among

**TABLE 2 |** Test of homogeneity of variances for scores on pre-test.

		Levene's statistic	df1	df2	Sig.
NAWLT	Based on mean	0.363	2	83	0.697
	Based on median	0.285	2	83	0.752
	Based on median and with adjusted df	0.285	2	80.090	0.753
	Based on trimmed mean	0.350	2	83	0.706
VKT	Based on mean	0.649	2	83	0.525
	Based on median	0.522	2	83	0.595
	Based on median and with adjusted df	0.522	2	77.808	0.595
	Based on trimmed mean	0.629	2	83	0.536

**TABLE 3 |** One-way between-groups ANOVA for the scores on pre-test.

ANOVA		Sum of squares	df	Mean square	F	Sig.
NAWLT	Between groups	6.058	2	3.029	0.674	0.512
	Within groups	372.791	83	4.491		
	Total	378.849	85			
VKT	Between groups	16.043	2	8.021	0.980	0.380
	Within groups	679.259	83	8.184		
	Total	695.302	85			

**TABLE 4 |** Descriptive statistics for post-test results.

		N	Mean	Std. deviation	Std. error	95% confidence interval for mean	
						Lower bound	Upper bound
NAWLT	EXP 1	31	20.42	3.686	0.662	19.07	21.77
	EXP 2	30	16.20	3.156	0.576	15.02	17.38
	CON	25	13.12	3.004	0.601	11.88	14.36
	Total	86	16.83	4.430	0.478	15.88	17.78
VKT	EXP 1	31	41.00	4.754	0.854	39.26	42.74
	EXP 2	30	34.73	5.889	1.075	32.53	36.93
	CON	25	30.92	6.474	1.295	28.25	33.59
	Total	86	35.88	6.993	0.754	34.38	37.38

a group of Psychology major university students and (b) comparing the learning outcomes from DFs to traditional materials. Designed as a semester-long experimental study, the results indicated that using DFs inherently could enhance students' engagement with learning their discipline-specific vocabulary items during the intervention. The findings also indicated that participants using DFs on their mobile devices outperformed the participants using paper flashcards and word lists in vocabulary learning. Moreover, the long-term impact of the interventions could be noteworthy with regard to sustaining

students' vocabulary retention rate at a higher level compared to the pre-treatment levels. The findings from the study generally add to the growing literature on mobile-assisted language learning by providing empirical support on the effectiveness of mobile-assisted vocabulary learning among students who have limited and discipline-specific exposure to English and minimal opportunity to utilize the gained knowledge outside the given academic context. In other words, these devices and technologies would act as facilitators and scaffolds in directing students' learning and enhancing their autonomy in taking control of their learning and hence practicing their agency in such a process.

The current study has implications for teaching academic vocabulary to university students. As English has established itself as the *lingua franca* for academic publication in international journals, university students are increasingly required to read and publish in English. In this regard, English has become much more instrumental in shaping their academic identity that also facilitates their access to the pertinent literature in their discipline (Paquot, 2010). Nevertheless, studies in academic writing indicated that English users with non-English linguistic backgrounds in particular face consequential linguistic impediments in such undertakings (Flowerdew, 2015), and inadequate vocabulary knowledge is one of the most important factors that add

**TABLE 5 |** Test of homogeneity of variances for scores on post-test.

		Levene's statistic	df1	df2	Sig.
NAWLT	Based on mean	1.636	2	83	0.201
	Based on median	1.333	2	83	0.269
	Based on median and with adjusted df	1.333	2	79.169	0.270
	Based on trimmed mean	1.633	2	83	0.202
VKT	Based on mean	1.956	2	83	0.148
	Based on median	1.547	2	83	0.219
	Based on median and with adjusted df	1.547	2	76.836	0.220
	Based on trimmed mean	1.923	2	83	0.153

**TABLE 6 |** One-way between-groups ANOVA for the scores on post-test.

#### ANOVA

		Sum of squares	df	Mean square	F	Sig.
NAWLT	Between groups	755.395	2	377.698	34.337	0.000
	Within groups	912.988	83	11.000		
	Total	1668.384	85			
VKT	Between groups	1467.131	2	733.565	22.637	0.000
	Within groups	2689.707	83	32.406		
	Total	4156.837	85			

**TABLE 7 |** Multiple comparisons.

#### Tukey's HSD

Dependent variable	(I) Group	(J) Group	Mean difference (I-J)	Std. error	Sig.	95% confidence interval	
						Lower bound	Upper bound
NAWLT	EXP 1	EXP 2	4.219*	0.849	0.000	2.19	6.25
		CON	7.299*	0.892	0.000	5.17	9.43
	EXP 2	EXP 1	-4.219*	0.849	0.000	-6.25	-2.19
		CON	3.080*	0.898	0.003	0.94	5.22
	CON	EXP 1	-7.299*	0.892	0.000	-9.43	-5.17
		EXP 2	-3.080*	0.898	0.003	-5.22	-0.94
VKT2	EXP 1	EXP 2	6.267*	1.458	0.000	2.79	9.75
		CON	10.080*	1.530	0.000	6.43	13.73
	EXP 2	EXP 1	-6.267*	1.458	0.000	-9.75	-2.79
		CON	3.813*	1.542	0.040	0.13	7.49
	CON	EXP 1	-10.080*	1.530	0.000	-13.73	-6.43
		EXP 2	-3.813*	1.542	0.040	-7.49	-0.13

\*The mean difference is significant at the 0.05 level.

to their disadvantage in academic publishing (Bazerman et al., 2012). Given the significant role of academic vocabulary in university studies, mastery over core academic words benefits university students and EAP programs in many ways. The findings of the current study indicated that using DFs and mobile applications provided the participants with the affordances to learn a considerable number of academic words frequently used in their disciplines. Accordingly, instructors and materials developers might consider adding them to EAP programs. Additionally, university students can use well-designed DF applications for developing their academic vocabulary knowledge in self-directed learning.

The study, however, has some limitations. First, the research was conducted with two relatively small experimental groups each containing around 30 participants and a control group with 25 individuals. This should be accounted for in generalizing the findings, as the small sample size in each group might have resulted in biased results. Second, given that the study lasted for a semester, controlling the students' possible exposure and contact with other language learning materials was not possible. In this regard, although they had no considerable exposure to other materials for learning academic vocabulary beyond the classroom, some learning might have resulted from other resources (Xodabande, 2018). Additionally, as vocabulary knowledge has different aspects and entails both receptive and productive uses of words (Nation, 2013), the current study was concerned with developing the participants' receptive knowledge of academic words only. This focus on receptive knowledge was in line with the participants' vocabulary learning needs (i.e., to read specialized and academic texts); however, it should be acknowledged that any intervention to develop

university students' productive vocabulary knowledge is much more important (Soyoof et al., 2022). Despite these limitations, the study was conducted in a longer time span with different measurements administered prior to and after the treatment, and the findings provided additional empirical evidence supporting the affordances of DFs and mobile devices for academic vocabulary learning in the EFL context. Future research might consider investigating the impacts of DFs not only on receptive knowledge of academic words but also on their productive use in speaking and writing.

## DATA AVAILABILITY STATEMENT

The datasets used and/or analyzed during the current study are available from the corresponding author on reasonable request.

## ETHICS STATEMENT

Ethical review and approval was not required for the study on human participants in accordance with the local legislation and institutional requirements. The patients/participants provided their written informed consent to participate in this study.

## AUTHOR CONTRIBUTIONS

All authors listed have made a substantial, direct, and intellectual contribution to the work and approved it for publication.

## REFERENCES

- Ashcroft, R. J., Cvitkovic, R., and Praver, M. (2018). Digital flashcard L2 vocabulary learning out-performs traditional flashcards at lower proficiency levels: a mixed-methods study of 139 Japanese university students. *EuroCALL Rev.* 26:14. doi: 10.4995/eurocall.2018.7881
- Atai, M. R., and Hejazi, S. Y. (2019). Assessment of academic English language needs of Iranian post-graduate students of psychology. *Iberica: Revista de La Asociación Europea de Lenguas Para Fines Específicos (AELFE)* 38, 275–302.
- Bazerman, C., Keranen, N., and Prudencio, F. E. (2012). "Facilitated immersion at a distance in second language scientific writing" in *University Writing: Selves and Texts in Academic Societies*. eds. M. Castelló and C. Donahue (Leiden: Brill), 235–248.
- Brown, D., Stoeckel, T., Mclean, S., and Stewart, J. (2020). The most appropriate lexical unit for L2 vocabulary research and pedagogy: a brief review of the evidence. *Appl. Linguis.* doi: 10.1093/applin/amaa061
- Browne, C., Culligan, B., and Phillips, J. (2013). The new academic word list. Available at: <http://www.newgeneralservicelist.org/nawl-new-academic-word-list> (Accessed May 17, 2022).
- Burston, J., and Giannakou, K. (2021). MALL language learning outcomes: a comprehensive meta-analysis 1994–2019. *ReCALL* 34, 147–168. doi: 10.1017/S0958344021000240
- Chen, Z., Chen, W., Jia, J., and An, H. (2020). The effects of using mobile devices on language learning: a meta-analysis. *Educ. Technol. Res. Dev.* 68, 1769–1789. doi: 10.1007/s11423-020-09801-5
- Chwo, G. S. M., Marek, M. W., and Wu, W.-C. V. (2018). Meta-analysis of MALL research and design. *System* 74, 62–72. doi: 10.1016/j.system.2018.02.009
- Cohen, J. (1988). *Statistical Power Analysis for the Behavioral Sciences*. London: Routledge.
- Coxhead, A. (2000). A new academic word list. *TESOL Q.* 34, 213–238. doi: 10.2307/3587951
- Coxhead, A. (2011). The academic word list 10 years on: research and teaching implications. *TESOL Q.* 45, 355–362. doi: 10.5054/tq.2011.254528
- Coxhead, A. (2018). *Vocabulary and English for Specific Purposes Research: Quantitative and Qualitative Perspectives*. London: Routledge.
- Coxhead, A. (2019). "Academic vocabulary," in *The Routledge Handbook of Vocabulary Studies*. ed. S. Webb (London: Routledge), 97–110.
- Coxhead, A., and Byrd, P. (2007). Preparing writing teachers to teach the vocabulary and grammar of academic prose. *J. Second. Lang. Writ.* 16, 129–147. doi: 10.1016/j.jslw.2007.07.002
- Dashtestani, R. (2016). Moving bravely towards mobile learning: Iranian students' use of mobile devices for learning English as a foreign language. *Comput. Assist. Lang. Learn.* 29, 815–832. doi: 10.1080/09588221.2015.1069360
- Dashtestani, R., and Hojatpanah, S. (2021). "Mobile-assisted language learning in a secondary school in Iran: discrepancy between the stakeholders' needs and the status quo," in *Handbook for Online Learning Contexts: Digital, Mobile and Open: Policy and Practice*. eds. A. Marcus-Quinn and T. Hourigan (Berlin: Springer International Publishing), 157–174.
- Dizon, G. (2016). Quizlet in the EFL classroom: enhancing academic vocabulary acquisition of Japanese university students. *Teach. Engl. Technol.* 16, 40–56.
- EFL Technologies (2017). NAWL builder multilingual (1.1). Available at: <https://apps.apple.com/us/app/nawl-builder-multilingual/id1097693935> (Accessed May 17, 2022).
- Evans, S., and Morrison, B. (2010). The first term at university: implications for EAP. *ELT J.* 65, 387–397. doi: 10.1093/elt/ccq072
- Evans, S., and Morrison, B. (2011). Meeting the challenges of English-medium higher education: the first-year experience in Hong Kong. *Engl. Specif. Purp.* 30, 198–208. doi: 10.1016/j.esp.2011.01.001



- Flowerdew, J. (2015). Some thoughts on English for research publication purposes (ERPP) and related issues. *Lang. Teach.* 48, 250–262. doi: 10.1017/S0261444812000523
- Frey, B. B. (2018). *The SAGE Encyclopedia of Educational Research, Measurement, and Evaluation*. Thousand Oaks, CA: SAGE Publications.
- Gardner, D., and Davies, M. (2014). A new academic vocabulary list. *Appl. Linguis.* 35, 305–327. doi: 10.1093/applin/amt015
- Honzard, R., and Soyoo, A. (2020). Vocabulary learning and retention: a comparison between a serious game and mobile application. *Int. J. Pedagogies Learn.* 15, 81–100.
- Hwang, G.-J., and Fu, Q.-K. (2019). Trends in the research design and application of mobile language learning: a review of 2007–2016 publications in selected SSCI journals. *Interact. Learn. Environ.* 27, 567–581. doi: 10.1080/10494820.2018.1486861
- Hyland, K., and Tse, P. (2007). Is there an “academic vocabulary”? *TESOL Q.* 41, 235–253. doi: 10.1002/j.1545-7249.2007.tb00058.x
- Kohnke, L., Zhang, R., and Zou, D. (2019). Using mobile vocabulary learning apps as aids to knowledge retention: business vocabulary acquisition. *J. Asia TEFL* 16, 683–690. doi: 10.18823/asiatfl.2019.16.2.16.683
- Kohnke, L., Zou, D., and Zhang, R. (2020). Exploring discipline-specific vocabulary retention in L2 through app design: implications for higher education students. *REL C J.* 52, 539–556. doi: 10.1177/0033688219899740
- Kornell, N. (2009). Optimising learning using flashcards: spacing is more effective than cramming. *Appl. Cogn. Psychol.* 23, 1297–1317. doi: 10.1002/acp.1537
- Laufer, B. (1996). “The lexical plight in second language reading: words you don’t know, words you think you know, and words you can’t guess,” in *Second Language Vocabulary Acquisition: A Rationale for Pedagogy*. eds. J. Coady and T. Huckin (Cambridge: Cambridge University Press), 20–34.
- Laufer, B. (2005). “Focus on form in second language vocabulary learning,” in *EUROSLA Yearbook: Volume 5*. eds. S. H. Foster-Cohen, P. G. M. del Mayo and J. Cenoz (Amsterdam: John Benjamins Publishing Company), 223–250.
- Laufer, B. (2013). Lexical thresholds for reading comprehension: what they are and how they can be used for teaching purposes. *TESOL Q.* 47, 867–872. doi: 10.1002/tesq.140
- Li, Y., and Hafner, C. A. (2022). Mobile-assisted vocabulary learning: investigating receptive and productive vocabulary knowledge of Chinese EFL learners. *ReCALL* 34, 66–80. doi: 10.1017/S0958344021000161
- Lin, J.-J., and Lin, H. (2019). Mobile-assisted ESL/EFL vocabulary learning: a systematic review and meta-analysis. *Comput. Assist. Lang. Learn.* 32, 878–919. doi: 10.1080/09588221.2018.1541359
- Liu, J., and Han, L. (2015). A corpus-based environmental academic word list building and its validity test. *Engl. Specif. Purp.* 39, 1–11. doi: 10.1016/j.esp.2015.03.001
- Ma, Q. (2017). “Technologies for teaching and learning L2 vocabulary,” in *The Handbook of Technology and Second Language Teaching and Learning*. eds. C. A. Chapelle and S. Sauro (Hoboken, NJ: Wiley-Blackwell), 45–61.
- Ma, Q., and Mei, F. (2021). Review of corpus tools for vocabulary teaching and learning. *J. China Comput.-Assisted Lang. Learn.* 1, 177–190. doi: 10.1515/jccall-2021-2008
- Mahdi, H. S. (2017). Effectiveness of mobile devices on vocabulary learning: a meta-analysis. *J. Educ. Comput. Res.* 56, 134–154. doi: 10.1177/0735633117698826
- Mansouri, B., and Mantero, M. (2019). “Knowledge construction in concept-based language instruction,” in *The TESOL encyclopedia of English language teaching*. eds. J. I. Lontos and M. DelliCarpini, (New Jersey: John Wiley & Sons, Inc.) 1–8.
- Masrai, A., and Milton, J. (2018). Measuring the contribution of academic and general vocabulary knowledge to learners’ academic achievement. *J. Engl. Acad. Purp.* 31, 44–57. doi: 10.1016/j.jeap.2017.12.006
- McLean, S., and Kramer, B. (2015). The creation of a new vocabulary levels test. *Shiken* 19, 1–11.
- Nagy, W., and Townsend, D. (2012). Words as tools: learning academic vocabulary as language acquisition. *Read. Res. Q.* 47, 91–108. doi: 10.1002/RRQ.011
- Nakata, T. (2019). “Learning words with flash cards and word cards,” in *The Routledge Handbook of Vocabulary Studies*. ed. S. Webb (London: Routledge), 304–319.
- Nation, I. S. P. (2013). *Learning Vocabulary in Another Language*. 2nd Edn. Cambridge: Cambridge University Press.
- Nazari, M., and Xodabande, I. (2020). L2 teachers’ mobile-related beliefs and practices: contributions of a professional development initiative. *Comput. Assist. Lang. Learn.* 1–30. doi: 10.1080/09588221.2020.1799825
- Nazari, M., and Xodabande, I. (2021). Discursive change in second language teachers’ online interactions: A microgenetic analysis. *Iranian journal of. Lang. Teach. Res.* 9, 63–76. doi: 10.30466/ijltr.2021.121076
- Paquot, M. (2010). *Academic Vocabulary in Learner Writing: From Extraction to Analysis*. London: Continuum International Publishing Group.
- Rassaei, E. (2020). Effects of mobile-mediated dynamic and nondynamic glosses on L2 vocabulary learning: a sociocultural perspective. *Mod. Lang. J.* 104, 284–303. doi: 10.1111/modl.12629
- Rassaei, E. (2021). Implementing mobile-mediated dynamic assessment for teaching request forms to EFL learners. *Comput. Assist. Lang. Learn.* 1–31. doi: 10.1080/09588221.2021.1912105
- Rasti-Behbahani, A., and Shahbazi, M. (2020). Investigating the effectiveness of a digital game-based task on the acquisition of word knowledge. *Comput. Assist. Lang. Learn.* 1–25. doi: 10.1080/09588221.2020.1846567
- Schmitt, N., Dunn, K., O’Sullivan, B., Anthony, L., and Kremmel, B. (2021). Introducing knowledge-based vocabulary lists (KVL). *TESOL J.* 12, 1–10. doi: 10.1002/tesj.622
- Schmitt, N., Jiang, X., and Grabe, W. (2011). The percentage of words known in a text and Reading comprehension. *Mod. Lang. J.* 95, 26–43. doi: 10.1111/j.1540-4781.2011.01146.x
- Seibert Hanson, A. E., and Brown, C. M. (2020). Enhancing L2 learning through a mobile assisted spaced-repetition tool: an effective but bitter pill? *Comput. Assist. Lang. Learn.* 33, 133–155. doi: 10.1080/09588221.2018.1552975
- Soyoo, A., Reynolds, B. L., Shadiev, R., and Vazquez-Calvo, B. (2022). A mixed-methods study of the incidental acquisition of foreign language vocabulary and healthcare knowledge through serious game play. *Comput. Assist. Lang. Learn.* 1–34. doi: 10.1080/09588221.2021.2021242
- Soyoo, A., Reynolds, B. L., Vazquez-Calvo, B., and McLay, K. (2021). Informal digital learning of English (IDLE): a scoping review of what has been done and a look towards what is to come. *Comput. Assist. Lang. Learn.* 1–27. doi: 10.1080/09588221.2021.1936562
- Spencer, S., Clegg, J., Lowe, H., and Stackhouse, J. (2017). Increasing adolescents’ depth of understanding of cross-curriculum words: an intervention study. *Int. J. Lang. Commun. Disord.* 52, 652–668. doi: 10.1111/1460-6984.12309
- Stockwell, G. (2013). “Technology and motivation in English-language teaching and learning,” in *International Perspectives on Motivation*. ed. E. Ushioda (London: Palgrave Macmillan), 156–175.
- Stoeckel, T., and Bennett, P. (2020). The New Academic Word List Test (NAWLTL). Available at: <https://www.newgeneralservicelist.org/ngsl-levels-test>. (Accessed March 22, 2020).
- Test your English (2022). Cambridge University Press & Assessment. Available at: <https://www.cambridgeenglish.org/test-your-english/general-english/> (Accessed May 17, 2022).
- Valizadeh, M., and Xodabande, I. (2021). General service and academic words in psychology research articles: a corpus based study. *J. AsiaTEFL* 18, 745–1070. doi: 10.18823/asiatfl.2021.18.3.15.959
- Webb, S., and Chang, A. C.-S. (2012). Second language vocabulary growth. *REL C J.* 43, 113–126. doi: 10.1177/0033688212439367
- Webb, S., and Nation, I. S. P. (2017). *How Vocabulary is Learned*. Oxford: Oxford University Press.
- Webb, S., Yanagisawa, A., and Uchihara, T. (2020). How effective are intentional vocabulary-learning activities? A meta-analysis. *Modern Lang. J.* 104, 715–738. doi: 10.1111/modl.12671
- West, M. (1953). *A General Service List of English Words*. London: Longman, Green & Co.
- Xodabande, I. (2017). The effectiveness of social media network telegram in teaching English language pronunciation to Iranian EFL learners. *Cogent Educ.* 4:1347081. doi: 10.1080/2331186X.2017.1347081
- Xodabande, I. (2018). Iranian EFL learners’ preferences of different digital technologies for language learning beyond the classroom. *Int. J. Educ. Literacy Stud.* 6, 20–31. doi: 10.7575/aiac.ijels.v6n.3p.20

- Xodabande, I., and Atai, M. R. (2020). Using mobile applications for self-directed learning of academic vocabulary among university students. *Open Learn J. Open Distance e-Learning*, 1–18. doi: 10.1080/02680513.2020.1847061
- Xodabande, I., Pourhassan, A., and Valizadeh, M. (2022). Self-directed learning of core vocabulary in English by EFL learners: comparing the outcomes from paper and mobile application flashcards. *J. Comput. Educ.* 9, 93–111. doi: 10.1007/S40692-021-00197-6
- Yang, L., and Coxhead, A. (2020). A corpus-based study of vocabulary in the new concept English textbook series. *RELC J.* doi: 10.1177/0033688220964162 [Epub ahead of print]
- Yüksel, H. G., Mercanoğlu, H. G., and Yılmaz, M. B. (2020). Digital flashcards vs. wordlists for learning technical vocabulary. *Comput. Assist. Lang. Learn.*, 1–17. doi: 10.1080/09588221.2020.1854312
- Zhang, R., Zou, D., and Xie, H. (2021). Spaced repetition for authentic mobile-assisted word learning: nature, learner perceptions, and factors leading to positive perceptions. *Comput. Assist. Lang. Learn.* 1–34. doi: 10.1080/09588221.2021.1888752

**Conflict of Interest:** The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

**Publisher's Note:** All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Copyright © 2022 Xodabande, Iravi, Mansouri and Matinparsa. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.



# The Impact of Mobile-Assisted Language Learning on English as a Foreign Language Learners' Vocabulary Learning Attitudes and Self-Regulatory Capacity

Xiao Lei<sup>1</sup>, Jalil Fathi<sup>2\*</sup>, Shabnam Noorbakhsh<sup>3</sup> and Masoud Rahimi<sup>2</sup>

<sup>1</sup>Department of Foreign Languages, Zunyi Medical University, Zunyi, China, <sup>2</sup>Department of English and Linguistics, Faculty of Language and Literature, University of Kurdistan, Sanandaj, Iran, <sup>3</sup>Department of English and Linguistics, Faculty of Language and Literature, Islamic Azad University, Sanandaj Branch, Sanandaj, Iran

## OPEN ACCESS

### Edited by:

Ehsan Rassaei,  
Majan University College, Oman

### Reviewed by:

Ali Soyoof,  
Monash University, Australia  
Javad Zare,  
Kosar University of Bojnord, Iran

### \*Correspondence:

Jalil Fathi  
jfathi13@yahoo.com

### Specialty section:

This article was submitted to  
Language Sciences,  
a section of the journal  
Frontiers in Psychology

**Received:** 10 February 2022

**Accepted:** 26 April 2022

**Published:** 21 June 2022

### Citation:

Lei X, Fathi J, Noorbakhsh S and  
Rahimi M (2022) The Impact of  
Mobile-Assisted Language Learning  
on English as a Foreign Language  
Learners' Vocabulary Learning  
Attitudes and Self-Regulatory  
Capacity.  
Front. Psychol. 13:872922.  
doi: 10.3389/fpsyg.2022.872922

Over the past decades, English as a foreign language (EFL) learning has witnessed a heightened interest in the role of mobile-assisted language learning (MALL) in vocabulary learning. To shed more light on the impact of MALL on vocabulary learning, this study, employing a quantitative longitudinal design, aimed at examining the impact of a MALL programme on 139 EFL learners' vocabulary learning attitudes and self-regulatory capacity. To this end, this study investigated the latent change score models of the learners' vocabulary learning attitudes and self-regulatory capacity over time. Over the course of 1 year, various mobile applications were integrated into the regular English language instruction of the learners. The required data were collected *via* administering vocabulary learning attitude and self-regulating capacity in vocabulary learning scales. The data were analysed applying latent growth curve modelling to examine the participants' longitudinal trajectories and patterns of change in the two waves of collected data. The fit indices of the latent change models revealed an increase in both the EFL learners' vocabulary learning attitudes and their self-regulatory capacity during the 1-year MALL programme. The analysis of between-person differences also indicated that changes in both variables were positively correlated.

**Keywords:** mobile-assisted language learning, vocabulary learning attitudes, self-regulatory capacity, EFL, latent curve modelling

## INTRODUCTION

As one of the viable and alluring technological devices, mobile devices have become widely available and advocated in educational settings (Looi et al., 2010; Wang et al., 2017; Hwang and Fu, 2018; Liu et al., 2018; Rassaei, 2019; Fathi et al., 2021). Mobile learning refers to the use of mobile technologies, such as smartphones and tablets (Huang et al., 2016), for educational purposes. Mobile-based learning is considered "learning across multiple contexts, through social and content interactions, using personal electronic devices" (Crompton, 2013, p. 4). In contrast

to conventional computer-based learning technologies, which are constrained by time and place, mobile-based technologies provide their users with ubiquitous access that is not constrained by spatial and temporal factors (Kukulska-Hulme and Traxler, 2005; Melhuish and Fallon, 2010; Wong, 2012; Wang and Shih, 2015). That is, the mobility of mobile-based technologies, like mobile phones, iPod, and tablets, provide second language (L2) learners with a convenient place and time and great learning opportunities (Sharples, 2006; Laurillard, 2007), which might not be supplied by desktop computers (Chinnery, 2006).

In line with mainstream education, the use of mobiles in language learning, or what is currently known as mobile-assisted language learning (MALL), gained much popularity (Kiernan and Aizawa, 2004; Hwang and Fu, 2018), as it was included in the foreign language curriculum and offering new learning devices to the net generation (Oblinger and Oblinger, 2005). MALL has received much interest because of its authentic and contextual language learning experiences (Chinnery, 2006; Kukulska-Hulme, 2006; Shadiev et al., 2017; Cheng and Chen, 2022). Previous research on MALL has examined the effect of mobile-based technologies on EFL learners' language learning skills, such as speaking (Sun et al., 2017), writing (Eubanks et al., 2018), listening (Jia and Hew, 2019), and reading (Chang and Hsu, 2011), and language sub-skills, such as grammar (Chu et al., 2019), vocabulary (Çakmak and Erçetin, 2018), and pronunciation (Wongsuriya, 2020). In comparison with all the MALL-related studies, the effect of mobile devices on vocabulary learning has attracted the most attention (Burston, 2014; Duman et al., 2015).

Although the impact of mobile-based technologies on English language vocabulary learning has been explored extensively (Stockwell, 2010; Çetinkaya and Sütçü, 2018; Lin and Lin, 2019), no large-scale studies have been carried out so far to check the potential impact of MALL on vocabulary learning attitudes in an EFL context. Following Masgoret and Gardner (2003), vocabulary learning attitude could be conceptualised as learners' positive/negative reactions towards their vocabulary learning. Positive attitudes towards second/foreign language learning (in this study, positive attitudes towards vocabulary learning) enhance learners' second/foreign language learning motivation and help them easily accomplish their language learning tasks (Ellis, 2008; Cheung and Hew, 2009; Nation and Webb, 2011; Burston, 2015; Chen et al., 2020; Soyoo et al., 2021, 2022). Therefore, positive attitudes towards vocabulary learning can facilitate second/foreign language learners' vocabulary learning.

The present study, therefore, is aimed to examine the influence of a 1-year MALL programme on EFL learners' vocabulary learning attitudes. That is, a quantitative longitudinal research design is adopted in this study to investigate EFL learners' longitudinal trajectories and patterns of change in vocabulary learning. In addition, as the concept of self-regulation (Zimmerman and Schunk, 2008), which refers to learners' activating and sustaining planned self-generated thoughts, emotions, and activities to achieve their learning goals (Zimmerman, 2000, 2008), has received considerable attention in the MALL context (e.g., García Botero et al., 2021), especially

in mobile-based vocabulary learning context (e.g., Barth et al., 2020). EFL learners' self-regulatory vocabulary learning changes over the 1-year MALL programme is also examined in the present study. Moreover, the reciprocity and relationship between EFL learners' vocabulary learning attitudes and their self-regulatory vocabulary learning are checked in this study.

As Engeström's (1987) activity theory, which was adopted in the current study, explores learners' gradual movement from other-regulation (i.e., interactive activities among peers) to self-regulation (i.e., learners' autonomous functioning), we investigated the EFL learners' self-regulatory vocabulary learning to check if EFL learners reach their self-regulation in vocabulary learning. In addition, as collaborative activities among EFL learners occur infrequently in the current study on the one hand, and EFL learners rarely use mobile-based applications for their English language courses, exploring the EFL learners' vocabulary learning attitudes in this new learning environment could shed more light on the related literature. The findings of the study can add to the literature given that no quantitative longitudinal growth modelling analysis has been carried out examining learners' vocabulary learning attitudes and self-regulation in the EFL context. The findings of the present study can also propose practical implications to contribute to EFL learners' vocabulary learning attitudes and their self-regulatory vocabulary learning in a mobile-based vocabulary learning context.

## LITERATURE REVIEW

The theoretical framework of the present study follows Engeström's (1987) activity theory. This theory originates from Vygotsky's (1978) social constructivism in which learners' self-regulation (i.e., internalisation of different skills) precedes their other-regulation (i.e., interactions and mediations among peers). Vygotsky's (1978) conceptualisation of learning comprises three components of subject (e.g., EFL learners and teachers), object or the desired goals that the learners want to achieve by the activity (e.g., improvements in vocabulary learning), and the mediational tools or artefacts (e.g., online applications) that mediate between the subject and object (see, for instance; Ebadi and Rahimi, 2019; Fathi and Rahimi, 2020; Rahimi and Fathi, 2021). However, activity theory conceptualises learning as interactions among the subject, object, and artefact, rules or norms (e.g., using mobile devices), community to which learners share a common goal (e.g., class and institution), and division of labor which refers to learners' responsibilities in accomplishing an activity in the community (e.g., instructors teaching EFL vocabulary and learners learning EFL vocabulary; Engeström, 2001). Based on this theory, learning could be regarded as an activity system in which learners are mediated by the tools to construct knowledge. Overtime, learners are expected to internalise the knowledge and act autonomously without receiving mediations from the other capable individuals and achieve their self-regulation.

The conceptualisation of activity theory is similar to the Vygotskian constructivist theory used in Ebadi and Rahimi

(2017, 2018, 2019), Rassaei's (2020), and Rahimi and Fathi (2021) study, as students are involved in collaborative activities to provide and receive both explicit and implicit mediations. In such collaborative activities, students gradually provide more explicit mediations on their peers' tasks only if the peers are not able to address their language issues through more implicit mediations. Students who address their language issues through minimal or implicit peer mediations are believed to be closer to their self-regulation in learning. However, activity theory goes beyond the Vygotsky's social constructivist theory by considering not only the students, their peer mediations, and their language learning goals, but also the rules or norms that the students are required to consider, like applying special language learning tools, the context where the students are engaged in the collaborative language learning activities, such as classroom context, and the students' responsibilities to do the target activity, like learning EFL vocabulary. That is, activity theory considers a wider range of factors while specifically focusing on students' collaborative learning activities.

Activity theory fits well with MALL design and, in particular, with mobile-based vocabulary learning, since it can "capture the dynamic nature of activity systems and changing points of focus over time" (Levy and Stockwell, 2006, p. 119). On the other hand, as mobile-based learning is an activity that happens in the physical and social environments (i.e., interactions among learners and the community, rules, and norms of the activities) of learning, it corresponds to the principles of activity theory (Wali et al., 2008). Moreover, a number of studies have adopted activity theory in mobile-based language learning environment (e.g., Uden, 2007; Liaw et al., 2010).

Following this theory, the EFL learners in the present study are mediated by their peers and the researcher/instructor to accomplish the vocabulary-based activities in the mobile learning environment. The learners, considered the subjects in the activity theory, form the learning community, since they all follow a common goal (i.e., improving their vocabulary knowledge) by sharing the responsibilities among the members to accomplish vocabulary-based activities using mobile applications. In the current study, activity theory helped explore the interactions between learners and their social and cultural contexts comprehensively, identify the constraints in the EFL MALL context, and capture longitudinal changes (e.g., Ellis, 2015; Soyoo et al., 2021) in EFL vocabulary learning attitudes and self-regulatory capacity in vocabulary learning. Activity theory has not been adopted in similar studies; as a result, the current study is believed to shed light on the literature.

## Mobile-Assisted Vocabulary Learning

Vocabulary learning, considered a prominent factor in English language learning (Nation, 2001, 2020; Zou et al., 2021), might be facilitated through mobile-assisted technologies (Hulstijn and Laufer, 2001). Following Sung et al. (2016), the user-friendliness of mobile-based technologies allows L2 learners to spend their free time learning new vocabulary items. In addition, applying mobile-assisted technologies, learners can learn at their own pace (Norris et al., 2011; Hung et al., 2012). For instance, some mobile vocabulary learning applications

allow L2 learners to download different types of content for offline study, others supply context-appropriate words by activating Global Positioning System (GPS) to identify learners' locations (Godwin-Jones, 2011).

A substantial body of studies have employed mobile phones for L2 vocabulary learning (Levy and Kennedy, 2005; Stockwell, 2010; Lai, 2016; Xu and Peng, 2017; Çetinkaya and Sütçü, 2018; Rosell-Aguilar, 2018; Lin and Lin, 2019; Seibert Hanson and Brown, 2020). Studies in this regard have specifically focused on the effect of short messages (Kennedy and Levy, 2008), electronic dictionaries (Song and Fox, 2008) and flashcards (Başoğlu and Akdemir, 2010). Lai (2016), for instance, applied WhatsApp (a mobile instant messenger) to provide a mobile immersion context in order to explore Chinese EFL students' vocabulary learning through an experimental-control group design. A number of high-frequency English verbs were introduced to the participants each session and the students were supposed to get engaged in some relevant text-chat discussions. Although the findings revealed no significant difference between the mobile-based and control groups' vocabulary learning after 3 months of 1.5-h weekly meetings, significant correlations were observed between the students' chat frequency and vocabulary gains in the mobile-based group. Lai also indicated that students' mentality towards the interactive learning environment in smartphone might facilitate or hamper the effectiveness of the mobile immersion.

Lin and Yu (2017) examined Taiwanese EFL learners' vocabulary learning through mobile phones. The participants of the study were supposed to learn four groups of target words each of which presented either in text, text and picture, text and sound, or text, picture, and sound mode in four different weeks. Having conducted a vocabulary test and a questionnaire related to cognitive load at the end of the programme, the findings indicated the significant role of audio input in recalling new word meanings and reducing the cognitive load of learning the new words. The findings also revealed the EFL learners' positive attitudes towards the mobile assisted vocabulary learning programme.

Wu (2015) also investigated the impact of smartphone application on EFL learners' vocabulary learning through an experimental research design. In the experimental group, the students covered the material through some mobile-based applications, while in the control group, the students covered the same lessons in print. The findings indicated that the students in the mobile-supported class outperformed those in the control group on their English vocabulary learning. Thornton and Houser (2005) further explored the effect of a mobile-based technology classroom on Japanese EFL students' vocabulary learning through a survey. The results showed that the students using mobile devices received higher vocabulary learning marks in comparison with the students learning on websites and paper. The students in the mobile-based technology classroom also had positive attitudes towards learning English idioms through mobile devices.

Similarly, Li and Hafner (2022) examined EFL learners' vocabulary learning through mobile-assisted word cards and paper word cards. They applied an experimental-control group



research design to check the learners' word knowledge in two areas of receptive knowledge of form-meaning connection (i.e., comprehending words in listening and reading) and productive knowledge of collocations (i.e., producing words in speaking and writing; Schmitt, 2010). The findings indicated that both the mobile-assisted and conventional classes improved the learners' vocabulary knowledge; however, the mobile-assisted class outperformed the conventional class in this regard. In addition, Liu (2016) exploring the effect of instant message function of smartphones on EFL learners' vocabulary knowledge and vocabulary retention, found that instant message group improved their vocabulary knowledge and retention and outperformed their conventional group counterpart. The results of the delayed post-tests further indicated that mobile-assisted learning was effective in long term provided that learning strategies be applied effectively.

In a similar vein, Rachels and Rockinson-Szapkiw (2018) applied a Duolingo-based instructional game course to develop L2 learners' vocabulary knowledge in a gamified learning environment. In comparison with a conventional instruction, Duolingo-based instruction greatly enhanced the learners' motivation while it had subtle impact on the learners' vocabulary learning. Loewen et al. (2019) also revealed language learners' improvement on Duolingo at the end of a term. Loewen et al. further showed a positive relationship between the amount of time spent on Duolingo and learning achievements. The participating learners also had positive perceptions towards the flexibility and gamification aspects of Duolingo. Lu (2008) and Zhang et al. (2011), respectively, exploring Taiwanese and Chinese EFL learners' vocabulary learning through mobile-assisted technologies, found that mobile-supported groups' vocabulary learning was improved and these improvements outperformed those of the conventional group; however, in Zhang et al. (2011) study, the vocabulary learning improvements did not last in the delayed post-test.

Additionally, Wong and Looi (2010) explored EFL students' English language preposition learning through two MALL case studies. The students using mobile devices constructed English sentences through some newly acquired prepositions. The students were consequently engaged in online or classroom discussion to develop their understanding of the English language prepositions. Finally, Kurt and Bensen (2017) adopted Vine (i.e., a mobile application to record and share videos) to help EFL learners develop their vocabulary learning by speaking and spelling through the mobile devices and review their knowledge by the video clips. The findings revealed that Vine-based practice improved the learners' vocabulary knowledge and enhanced their motivation for learning vocabulary.

## Self-Regulation in Vocabulary Learning

Self-regulation could be conceptualised as "the process by which learners personally activate and sustain cognitions, affects and behaviours that are systematically oriented towards the attainment of learning goals" (Zimmerman and Schunk, 2008, p. vii). Having originated in Corno and Mandinach's (1983) theory, self-regulation comprises three main components of strategic actions, which comprises planning, monitoring, and evaluating

strategies, learning motivation, which focuses on learner engagement, and self-efficacy, which focuses on learner abilities to achieve a goal (Chen and Hsu, 2020).

In spite of a significant body of research pertaining to vocabulary learning strategies (Tseng and Schmitt, 2008; Chacón-Beltrán, 2018; Wyrá and Lawson, 2018; Schmitt and Schmitt, 2020; Teng, 2020), there have been various conceptualisations regarding the concept of learning strategy (Dörnyei, 2005; Tseng et al., 2006; Schmitt and Schmitt, 2020). As a different perspective towards vocabulary learning strategies, Tseng et al. (2006) regarded strategic vocabulary learning as learners' self-regulation in language learning and developed a questionnaire to examine learners' self-regulatory capacity in vocabulary learning. Tseng et al. questionnaire incorporates the concept of self-regulation into the field of L2 learning and operationalises learning strategies as self-regulatory capacity. In other words, it is argued that strategic learning has been reconceptualised through the theoretical lens of self-regulated learning. However, not many empirical studies have investigated the concept of self-regulating capacity in vocabulary learning, especially in EFL settings.

According to Azevedo and Cromley (2004), learners with high self-regulation skills outperform those with low self-regulation skills when it comes to technology-enhanced learning space. On the other hand, technology-enhanced learning tools facilitate such self-regulation skills as goal setting, task-based strategies, time management, collaborative learning activities, and self-evaluation (Carneiro et al., 2007; Shea and Bidjerano, 2010; Lai and Gu, 2011; Fathi et al., 2018; Zheng et al., 2018). Following Chen et al. (2008) and Sha et al. (2012), self-regulated learning skills and mobile-based technologies are closely interrelated and promoting either of them can heavily influence the other one. It is also argued that the flexibility of mobile-based technologies in delivering content and increasing motivation and engagement can positively affect students' self-regulation (Sung et al., 2015, 2016) which can consequently raise their self-awareness, self-study behaviour (Kondo et al., 2012), and positive attitudes (Zimmerman and Schunk, 2008; Ning and Downing, 2010).

The concept of self-regulation is an under-researched concept in the MALL context (Sha et al., 2012; Hernández and Rankin, 2015; García Botero et al., 2021). In the EFL context, Chen et al. (2019), for instance, explored Taiwanese EFL learners' vocabulary learning along with a self-regulation mechanism through a mobile-assisted application. Adopting an experimental-control group design, the learners in one group received an English vocabulary-based application with a self-regulation mechanism while the other group received English vocabulary learning application without a self-regulation mechanism. The findings showed that learners receiving self-regulated learning had better learning performance and motivation.

In a similar vein, Barth et al. (2020) exploring Chinese EFL students' attitudes and perceptions towards design factors supporting self-directed mobile-based vocabulary learning, found that scaffolding in the form of first language translation developed basic and high-frequency vocabulary learning in mobile learning environments. The findings further propose that in order to

move mobile learning to mainstream education, students' capacity for self-directed learning need to be enhanced beforehand. Haq (2019) also indicated that MALL positively influenced EFL learners' self-regulated learning strategies. Haq maintained that there were no differences between self-regulated learning strategies of EFL learners with different proficiency levels. Additionally, García Botero et al. (2021) exploring L2 students' self-regulation in a mobile-based environment using Duolingo, revealed that students being trained in self-regulation showed a higher participation in Duolingo and higher test marks in L2 writing.

## PURPOSE OF THE STUDY

As the literature review revealed, mobile-based applications had positive effects on EFL vocabulary learning attitudes (Wu, 2015; Kurt and Bensen, 2017; Lin and Yu, 2017) and self-regulatory capacity in EFL vocabulary learning (Chen et al., 2008; Sha et al., 2012; Haq, 2019). However, vocabulary learning attitudes and self-regulatory capacity in vocabulary learning have not been explored in quantitative longitudinal research designs in EFL context (Zhang and Zou, 2020). Given the fact that L2 learning takes a long time, comprehensive understanding of L2 processes and changes should be investigated over time (Ortega and Byrnes, 2009). Therefore, we employed a quantitative longitudinal research design and adopted Engeström's (1987) activity theory to examine the EFL learners' longitudinal trajectories and patterns of change for vocabulary learning attitudes and self-regulatory capacity in vocabulary learning in a 1-year MALL programme. The reciprocity and relationship between these two factors were also examined. Employing quantitative longitudinal research designs have "several conceptual, methodological, and practical advantages and can stimulate the development and empirical examination of more complex questions and models concerning L2 development over time" (Barkaoui, 2014, p. 65). To accomplish the purposes of the study, the following research questions are formulated:

1. What are the effects of a 1-year MALL programme on EFL learners' vocabulary learning attitudes and self-regulatory vocabulary learning?
2. What are the relationships between EFL learners' vocabulary learning attitudes and self-regulatory vocabulary learning?

## MATERIALS AND METHODS

### Design of the Study

The current study adopted a longitudinal research design to collect quantitative data about EFL learners' vocabulary learning attitudes and self-regulatory vocabulary learning in a 1-year MALL programme. Caruana et al. (2015) defined longitudinal studies as continuous or repeated investigations of learners over a long time without any external influences being applied. Longitudinal research is believed to provide the researchers with rich data, since it investigates the participants in a long period of time.

## Participants

The participants recruited for the purpose of the present study were 175 EFL learners (78 males, and 72 females) in the age range of 16–40, studying in a private English language institute in Kurdistan, Iran. Convenience sampling method (Dörnyei, 2007) was applied to choose the participating learners in the present study. That is, sample selection followed the researcher/instructor's certain practical criteria, like the participants' "geographical proximity," their "availability at a certain time," their "easy accessibility," and their "willingness to volunteer" in the current study (Dörnyei, 2007, p. 99). As the researcher was teaching at that private language institute, selecting the learners, who were studying at the language institute, was convenient for the researcher. However, at the beginning 4 months of the study, 36 participants dropped out of the study due to being reluctant and/or unmotivated to continue the intervention. Thus, 139 EFL learners, as the total number of participants of this quantitative longitudinal study, underwent a 1-year MALL programme.

The participants were not at the same English language proficiency level. Before embarking upon the study, all the learners had gone through 4–7 years of English learning course in public school; however, some of the participants had experienced more than 7 years due to the additional university education. As a result, all the learners took full-course placement test which was designed and validated by the authors. Based on the test results, the learners were assigned to beginner, intermediate, or advanced level to cover different coursebooks of Top Notch Fundamentals, 1, 2, 3, and Summit 1 and 2. Except for five participants, the rest ( $N=134$ ) had computer literacy and they all had access to smart phones. The information about the participants is provided in more detail in **Table 1**.

## Instruments

### Vocabulary Learning Attitude Scale

In order to measure the vocabulary learning attitudes of the participants in the current study, vocabulary learning attitude scale (VLAS) developed by Tseng and Schmitt (2008) was

**TABLE 1 |** Demographic information of the participants.

		Frequency	Percentage
Gender	Female	61	43.8
	Male	78	56.9
	Total	139	100.0
Age	16–25	103	74.1
	Above 25	36	25.8
	Total	139	100.0
Computer literacy	Yes	134	96.4
	No	5	3.5
	Total	139	100.0
Smart phone availability	Yes	139	100.0
	No	0	0.0
	Total	139	100.0
Multilingualism	Kurdish	137	98.5
	Farsi	0	0.0
	Kurdish-Farsi	2	1.4
	Total	139	100.0

employed. The scale comprises 10 items, each targeting the attitude of participants towards their vocabulary retrieval. It is a six-point Likert scale that range from 1 = “strongly agree” to 6 = “strongly disagree” in 10–60 score range. The validity of VLAS was found to be acceptable. Additionally, the internal consistency of the scale, as measured by Cronbach’s Alpha formula, was reported to be 0.86 in this study.

### Self-Regulating Capacity in Vocabulary Learning Scale

To measure the EFL learners’ self-regulatory capacity in vocabulary learning, self-regulating capacity in vocabulary learning scale (SRCV), which was developed and validated by Tseng et al. (2006), was employed in the present study. SRCV includes 20 items with six-point Likert scale which each ranges from 1 = “strongly agree” to 6 = “strongly disagree”; the scale is in 20–120 score range. This instrument addresses five subscales of commitment control (i.e., setting goals for vocabulary learning); metacognitive control (i.e., keeping concentration on vocabulary learning); satiation control (i.e., overcoming boredom in vocabulary learning); emotion control (i.e., controlling stress in vocabulary learning); and environment control (i.e., dealing with contextual factors effectively). The validity of SRCV was checked by Tseng et al. (2006) and the results were acceptable. In addition, the internal consistency of SRCV (i.e., the reliability of SRCV) was measured through Cronbach’s Alpha the present researchers and the outcome was 0.85, which was also acceptable.

### Procedure

The intervention started at the end of June, 2015 and lasted at the end of June, 2016. The main coursebooks were *Top Notch and Summit series* which were required to be covered during the 1-year intervention. To homogenise different groups of learners and place them at their right level of language proficiency, all the participants were required to take a full-course placement test. The results were interpreted following the standards of the language institutes as A1 for Top Notch Fundamentals and Top Notch 1, A2 for Top Notch 2, B1 for Top Notch 3, B2 for Summit 1, and C1 for Summit 2. The participants were then provided with one of Top Notch coursebooks appropriate to their proficiency level. Prior to the treatment, SRCV and VLAS were administered by the researcher/instructor to specify the learners’ initial level of self-regulation and attitudes towards EFL vocabulary learning. Next, a set of multiple mobile applications were integrated into the 1-year MALL programme. At the end of the 1-year MALL programme, SRCV and VLAS were re-administered to collect the second wave of the data.

### The Applications Used in the MALL Programme

The list of the language learning applications and their descriptions employed in the MALL programme is summarised in Table 2.

The selection of the mobile-based applications presented in Table 2 was due to their convenient accessibility, their free charge, the researcher/instructor’s familiarity and previous experience of using such language learning and mobile-based

applications, and popularity of these applications among the learners. The following sections explain and exemplify the mobile-based applications used in the current study.

### Merriam-Webster and Longman e-Dictionaries

First, the researcher/instructor decided to apply various dictionary applications including Longman Dictionary of Contemporary English and Merriam Webster Dictionary for the learners of all levels in order to allow them to use the dictionaries for different purposes. The researcher/instructor explained how to install the applications and how to look up the words in the dictionaries *via* their mobiles. That is, the learners were taught how to use the dictionaries for word definitions, pronunciations, part of speech, idioms, and collocations. The learners were also allowed to use these applications freely inside the class whenever they deemed them necessary. Upon applying the words accurately, the researcher/instructor could make sure that the learners regularly used the dictionaries in and outside the classroom. In addition, the researcher/instructor regularly monitored the learners while using the dictionaries.

**TABLE 2 |** The list of language learning applications and their descriptions.

Name of applications	Description
Merriam-Webster and Longman e-dictionaries	Provide various definitions of a term Provide collocations and idioms Provide British and American pronunciations Faster, easier, and better usage and environment
Widget of 504 and 1,100 Essential Words G5	Provides vocabularies used in sentences with Persian meaning Provides synonyms and antonyms Reminds automatically to review vocabularies daily
English songs	Improves listening with authentic songs Provides the meaning of the songs Evaluates learners’ understanding and listening with quizzes Attractive environment for better confirmation of covered vocabularies and structures
Social network applications	Provide an interactive atmosphere to speak Assimilate synchronous learning Assimilate asynchronous learning
Beetalk and call recorder	Games and Puzzles Offer a collaborative room to share ideas
Vocabulary for high school students	Allow learners to record their voice Provides a picture related to the meaning of vocabularies Provides pronunciation and synonyms and parts of speech of vocabularies Provides audio and video files
Encarta student premium	Gives a dictionary of definitions and thesauruses Provides lots of authentic materials

### **Widget of 504 and 1,100 Essential Words G5**

In order to improve the vocabulary knowledge of participants, the researcher/instructor also provided the learners with different Widgets of vocabulary expanding applications including *400 Essential Words* for A1 and A2 levels, *504 Absolutely Essential Words* for B1 and B2 levels, and *1,100 Words* for C1 level learners. By regularly monitoring and asking the learners about the target words, the researcher/instructor was assured that the learners regularly checked these vocabulary applications in and outside the classroom.

### **English Songs**

To make the learning more interesting, the learners at all levels were introduced to *English Songs* application, enjoying three different levels appropriate for elementary, intermediate, and advanced EFL learners. Each level includes 20 songs allowing learners to learn more vocabulary in the context of songs accompanied by having access to the lyrics with Persian translation provided under each line. The learners could also access to the list of vocabulary derived directly from the songs. The researcher/instructor frequently checked the learners' vocabulary knowledge.

### **Social Network Applications**

To make the courses more communicative, the researcher/instructor decided to provide the participants with social network applications, such as *Telegram* and *WhatsApp*. In other words, the main purpose of such applications was to make the EFL learners much more engaged in their vocabulary learning. Therefore, the researcher/instructor administered some groups in Telegram through which the learners could chat together. The learners were involved in different discussions by asking and answering questions. The topics were specified either by the researcher/instructor or by the learners themselves. Taking part in English discussions in these websites was welcomed and encouraged. The group members were told to feel free to send their messages, ask their questions, post pictures, send English quotations, and even share *Word of the Day* from their e-dictionary applications. The researcher/instructor was the manager and one of the members of these groups; as a result, different activities of the learners were monitored and checked.

### **Beetalk and Call Recorder**

The learners were also exposed to *Beetalk*, which is a social network application, as well as a *Call Recorder* to record their voice while they had to make a phone conversation. They were all supposed to conduct a phone conversation according to the samples provided in their books and share the recorded files through *Beetalk* to get feedback from the researcher/instructor and their classmates. Each learner made a number of phone conversations, recorded them, and then shared the recorded voices with other learners through *Beetalk* which allowed the learners to have more discussions about the recorded voices, especially about the use of vocabulary.

Afterwards, the learners were provided with MP3 audio files and audiobooks. There were numerous benefits for mobile phone-based audiobooks in comparison with their CD-ROM or cassette counterparts. They displayed a rich auditory context for the learners with different proficiency levels and in various contexts. The learners were encouraged to listen to the audio files through their mobile phones in different situations. Since they all had access to the pdf files of all stories on the website, they could enjoy using them whenever they found the MP3 stories unclear. For the purpose of the present study, the researcher/instructor introduced the learners to four-minute mp3 stories downloaded from [www.americanstoryteller.com](http://www.americanstoryteller.com). By regularly asking the learners about the vocabulary and content of the stories and having further discussions in the classroom, the researcher/instructor ensured that the learners listened to the audios outside the classroom and understood the vocabulary and content as well.

### **Vocabulary for High School Students**

As a part of the programme, *Vocabulary for High School Students* application was also provided to all the participants. It is rich in terms of having various features and providing an appropriate atmosphere for vocabulary learning. It facilitates learning vocabulary by main ideas, Anglo-Saxon, Latin, and Greek prefixes, roots, and other word elements and providing the EFL learners with accurate pronunciation, English definition and synonyms, and parts of speech, all of which were presented in either pictorial or auditory mode. The researcher/instructor made sure that the learners regularly checked *Vocabulary for High School Students* application in and outside the classroom through monitoring and asking the learners about the words.

### **Encarta Student Premium**

Finally, to enrich the learners' vocabulary knowledge by presenting authentic materials, the researcher/instructor decided to install *Encarta Microsoft Student Encyclopedia*, which includes templates and tutorials to help learners do their assignments in Microsoft Office. Microsoft Office also includes *Encarta Premium* which provides encyclopaedia articles, photos and illustrations, videos and animations, audios, maps, and website addresses. *Encarta Dictionary* is another useful feature of *Encarta Premium* as it provides thesaurus to find synonyms and antonyms, translations to translate a word or phrase into another language, as well as verb conjugation to conjugate verbs. Each session, the researcher/instructor specified different tasks for learners to accomplish using *Encarta Microsoft Student Encyclopedia*. In the following session, the researcher/instructor asked the learners about different vocabulary and content of the application to substantiate that the learners applied the application properly.

## **Data Analysis**

Following McArdle (2001) and Ferrer and McArdle (2010), we employed latent change score models and used the programme Mplus 7.11, introduced by Muthén and Muthén (2012), to measure changes. Based on this model, change is conceptualised as defining the variables at the second time wave and compare



them with the sum of the score at the first time wave (McArdle and Prindle, 2008). Based on this model, change in one variable can be related to change in another variable (McArdle and Hamagami, 2001). To address the purpose of this study, we utilised the bivariate change models which include auto-regression paths in addition to crossed regression paths from the variables (i.e., vocabulary learning attitudes and self-regulatory capacity in vocabulary learning) at Time Wave 1 to the change score variables, the correlation of vocabulary learning attitudes and self-regulatory capacity in vocabulary learning at Time Wave 2, and the correlation of change scores.

To assess goodness-of-fit of the models and the  $\chi^2$  value, root mean square error of approximation (RMSEA), standardised root mean square residual (SRMR), and comparative fit index (CFI) were checked. Following Hu and Bentler (1999), the model is viewed as acceptable at  $CFI \geq 0.90$ . In addition, Hu and Bentler (1999) suggest that for a good fit, SRMR should not be more than 0.08. Furthermore, RMSEA values  $\leq 0.06$  is regarded to be of good fit, RMSEA value  $\leq 0.08$  is of fair fit, RMSEA value between 0.08 and 0.10 is considered mediocre fit, and RMSEA value  $> 0.10$  is regarded as poor fit (Hu and Bentler, 1999). **Table 3** shows the descriptive statistics and associations of the latent factors in vocabulary learning attitudes and self-regulation in EFL vocabulary learning.

As shown in **Table 3**, the variables were correlated significantly at different time waves. There were positive, medium-sized, and negative relationships between the variables. However, the relationship between vocabulary learning attitudes at Time Wave 1 and self-regulatory vocabulary learning at Time Wave 2 was not significant.

## RESULTS

### Within-Person Change

According to McArdle and Prindle (2008), factor structure and longitudinal invariance need to be tested prior to checking the latent change score models. Univariate latent change score models of EFL learners' vocabulary learning attitudes and their self-regulatory vocabulary learning were set up to examine the amount of change in the two variables over time.

In order to investigate vocabulary learning attitude change over time, a latent change score model was set in which the overall fit indices of the model turned out to be acceptable

( $\chi^2/df=1.79$ ,  $CFI=0.96$ ,  $RMSEA=0.05$ ,  $p<0.00$ ). In addition, the latent change variable mean was significant [ $M=0.28$  ( $SE=0.03$ ),  $p<0.00$ ] revealing that there was a significant increase of vocabulary learning attitudes from Time Wave 1 to Time Wave 2. Furthermore, vocabulary attitudes at Time Wave 1 with  $\sigma=0.15$  ( $SE=0.02$ ),  $p<0.00$  and the latent change variable with  $\sigma=0.19$  ( $SE=0.03$ ),  $p<0.00$  indicated a significant value of variance, reflecting individual differences both in vocabulary learning attitudes at Time Wave 1 and the change to Time Wave 2.

Concerning self-regulatory vocabulary learning change over time, the fit indices of the latent change score model for self-regulatory capacity in EFL vocabulary learning was acceptable ( $\chi^2/df=1.68$ ,  $CFI=0.94$ ,  $RMSEA=0.04$ ,  $p<0.00$ ). The latent change variable mean was significant [ $M=0.24$  ( $SE=0.03$ ),  $p<0.00$ ] indicating an increase of self-regulatory vocabulary learning from Time Wave 1 to Time Wave 2. Moreover, there was a significant variance at Time Wave 1 with  $\sigma=0.21$  ( $SE=0.04$ ),  $p<0.00$  and also a significant variance in the latent change variable with  $\sigma=0.19$  ( $SE=0.02$ ),  $p<0.00$ . Therefore, it can be suggested that there were significant individual differences within the variables under investigation.

### Between-Person Differences in Change

The two univariate latent change score models were integrated in one bivariate latent curve model to examine the dynamic relationships of EFL learners' attitudes towards vocabulary learning and their self-regulatory vocabulary learning over time. In this bivariate model, the predictive impact of both variables at a previous stage (i.e., Time Wave 1) in predicting change in the other variable can be tested. Moreover, through the bivariate model, it can be tested if there is a correlation between both change variables. The fit indices of the bivariate model were acceptable ( $\chi^2/df=1.98$ ,  $CFI=0.95$ ,  $RMSEA=0.04$ ,  $p<0.00$ ).

Moreover, the relationship of vocabulary learning attitudes and self-regulatory capacity in vocabulary learning at Time Wave 1 was significant and positive ( $\beta=0.44$  and  $p<0.00$ ). The association of change variables was also significant and positive ( $\beta=0.39$  and  $p<0.00$ ). This means that learners who demonstrated high improvements in vocabulary learning attitudes had also strong improvements in self-regulatory capacity in vocabulary learning. This indicated that high levels of vocabulary learning attitude at Time Wave 1 were related to high levels of self-regulatory capacity increases over time.

## DISCUSSION

Drawing on Engeström's (1987) activity theory, the current study was set to examine the impact of a MALL programme on EFL learners' vocabulary learning attitudes and self-regulatory capacity in vocabulary learning through a quantitative longitudinal research design. Two latent change score models and their interrelations were tested in the present study: within-person changes of vocabulary learning attitudes and self-regulatory vocabulary learning and between-person differences

**TABLE 3 |** Descriptive statistics and latent inter-correlations of variables.

	M (latent mean)	SD	1	2	3
Vocabulary attitude 1	2.23	0.39			
Vocabulary attitude 2	2.52	0.58	0.697		
Self-regulation 1	3.26	0.31	0.472	0.367	
Self-regulation 2	3.37	0.35	0.192	0.276	0.573

(individual differences). First, the results of within-person differences revealed overall increase in both vocabulary learning attitudes and self-regulatory vocabulary learning. In other words, the involvement of learners in MALL activities contributed to positive vocabulary learning attitudes and self-regulatory capacity in vocabulary learning over time.

The results of the present study could be in line with the findings of Ellis (2008), Burston (2015), and Liu (2016) who confirmed the beneficial role of mobile-integrated language learning procedures in creating positive learning attitudes. Rachels and Rockinson-Szapkiw (2018) also argued that mobile-based instruction developed learners' positive perceptions towards vocabulary learning. The findings of this study seem justifiable since mobile devices are likely to have provided the participants with collaborative activities both in and outside the classroom, thereby, contributing to foster positive learning attitudes among EFL learners. This means that the mobile applications possessed a lot of affordances for involving the learners in collaborative activities so that the learners could carry out the vocabulary tasks more conveniently and enthusiastically. In agreement with the findings, Lu (2008) also claims that MALL encourages interactions among the learners, enhances integrations among natural communication needs and language learning, and fosters retention of language learning skills.

Furthermore, following Engeström's (1987) activity theory, the researcher/instructor's and the other peers' mediations engaged the learners in vocabulary-based activities. The instructor's and other peers' mediations along with the MALL environment helped the learners reach their self-regulation. Different applications available in the MALL space helped the learners have more effective and efficient peer mediations on each other's communicative activities, which could further contributed to their vocabulary improvements. The findings of the present study showed that the MALL programme helped the EFL learners engage in other-regulation activities, following Engeström's (1987) activity theory, as they were willing to use the MALL applications for vocabulary improvement purposes. The learners' other regulations in the MALL environment also helped the learners reach their self-regulation in vocabulary development. That is, the learners gradually became autonomous in using the MALL applications appropriately for improving their vocabulary knowledge. This could be regarded as another evidence of the MALL learners' high vocabulary improvements. The EFL learners also had more positive vocabulary learning attitudes which might further substantiate the above-mentioned findings of the study.

With regard to the improvements in self-regulatory vocabulary learning, the findings of the present study are consistent with those of Shea and Bidjerano (2010) and Lai and Gu (2011) who emphasised the effects of technological devices on fostering self-regulated learning. According to Zimmerman (2000), self-regulated learners are able to become meta-cognitively, motivationally, and behaviourally active participants of their own learning processes. From this perspective, the learners of the present study were meta-cognitively, motivationally, and behaviourally active in the sense that MALL helped them set more appropriate goals

and use effective strategies to manage processes and content and to engage in help-seeking, self-monitoring, and self-evaluation. It could be argued that MALL marks a shift from teacher-driven learning to student-driven learning which might help students become autonomous and active learners, take the responsibility of their own learning, and be able to self-regulate their learning processes. In addition, as mobile vocabulary learning applications provide the language learners with the opportunity to "control their learning process, learn at their own pace and learn for their learning needs without being confined in fixed class schedules and physical settings" (Wang and Shih, 2015, p. 373), they enable learners to self-regulate their learning processes.

These findings are also consistent with Engeström's (1987) activity theory, as learners self-regulation of learning is the desired outcome which they try to achieve following their interactions with other peers. Based on the Engeström's (1987) activity theory, the MALL applications allowed the more capable peers to use self-regulatory strategies to mediate the less capable learners so that they could increase their vocabulary knowledge in the MALL environment more effectively. The community of more and less capable learners in the MALL environment allowed the less capable learners became more engaged in the communicative tasks as the capable peers were there to contribute to their vocabulary knowledge. These peer mediations between the less and more capable learners in the MALL space allowed both the more and less capable learners improve their vocabulary knowledge, and hence, their in self-regulatory vocabulary learning.

Between-person differences of change were also examined to address the second purpose of the present study. The interrelation of EFL vocabulary learning attitudes and self-regulatory capacity in vocabulary learning was investigated using bivariate change models in which each variable is considered a predictor for the other. The findings in this regard indicated that change in vocabulary learning attitudes and self-regulatory capacity in vocabulary learning was positively correlated. That is, EFL learners who showed high improvements in vocabulary learning attitudes were willing to show more self-regulatory capacity in vocabulary learning, whereas EFL learners who showed less improvements in vocabulary learning attitudes had a tendency to show less self-regulatory capacity in vocabulary learning. The reciprocal relationship between learning attitudes and self-regulated learning was corroborated by Zimmerman and Schunk (2008) and Ning and Downing (2010). These findings demonstrated that learners with positive learning attitudes were more likely to show self-regulatory and achievement-oriented behaviours. Based on activity theory (Engeström, 1987), learners who were actively engaged in vocabulary-based interactions in MALL environment (i.e., other-regulating their vocabulary knowledge) were better able to achieve their self-regulation of vocabulary learning.

The current study employed a dynamic rather than a static procedure to assess changes *via* latent change score models. This procedure helped identify model changes in

EFL learners' vocabulary learning attitudes and learners' self-regulatory vocabulary learning, as well as their interrelation, allowing the investigation of within-person changes and between-person differences in change. However, the EFL learners' vocabulary learning attitudes and self-regulatory vocabulary learning were measured through self-report scales. Moreover, since the changes in both variables were assessed simultaneously, causality inference in the relationship of the two variables changes were not possible. Given the fact that the present study utilised a two-wave set of data in the analysis, future EFL researchers are recommended to include more time waves and investigate more detailed analyses of change. Following Ellis (2015), in order to confirm changes in language acquisition, research data need to be based on learners' social activities or participations and learners' change in a particular and different context.

## CONCLUSION

The current study showed the importance of MALL in fostering EFL learners' vocabulary learning attitudes and self-regulatory capacity in vocabulary learning. With regard to the fruitful results of the present study, implications for EFL learners, teachers, and educators are suggested. For instance, the vital role of MALL in enhancing EFL learners' vocabulary learning attitudes and regulating their English vocabulary learning were confirmed in the present study, although some learners did not consider MALL efficient in regulating their vocabulary knowledge. The insular learners changed their mind about efficiency of MALL in boosting their vocabulary knowledge, as they found themselves enriched in pervasive domination of vocabulary at the end of the course due to using numerous mobile applications.

Not only language learners, but also EFL teachers can benefit from the findings of the present study. For instance, they may enjoy mobile-based applications to support the teaching process. To this aim, EFL teachers need to be equipped with a pre-packed collection of mobile applications appropriate for their language teaching and learning context. The results of this study further encourage EFL teachers not to see mobile phones as an interruptive device in EFL education contexts but rather as a great facilitator of English vocabulary learning in classroom. It is also implied that mobile-based applications facilitate both learner-learner and learner-teacher interactions. Çetinkaya and Sütçü (2018), in this regard, argue that mobile-based applications help learners have more in-class and out-of-class interactions. Moreover, Çetinkaya and Sütçü (2018) claim that mobile applications expedite the question-and-answer interplay between learners and teachers. Following the findings of the current study, EFL teacher educators and policy makers might provide and facilitate a convenient condition for EFL teachers and learners to apply such mobile-based applications in their vocabulary learning courses.

Additionally, incorporating mobile-based applications for improving the vocabulary knowledge of EFL learners, enhances

both learner-learner and instructor-learners interactions during and outside the class. EFL learners can have and receive more peer- and instructor- vocabulary mediations and improve their vocabulary knowledge in that regard (Rassaei, 2020). Incorporating mobile-based applications can also increase EFL learners' motivation, which can help them enhance their vocabulary knowledge and self-regulatory vocabulary learning. However, it is important to instruct EFL instructors to incorporate appropriate mobile-based applications in their EFL courses. EFL learners need to be familiarised with the mobile-based applications, so as to apply the applications not only when they are engaged in collaborative activities with their peers, but also when they use the applications autonomously (when they achieve their self-regulation in vocabulary learning), like those of Rahimi and Fathi (2021) and Rassaei's (2021) study in which learners achieved their self-regulation after providing and receiving implicit and explicit peer mediations. Following Simpson (2005), getting an acceptable level of digital literacy seems to be a fundamental prerequisite to conduct mobile-based courses in EFL contexts. Digital literacy could be defined as "the ability to use information and communication technologies to find, evaluate, create, and communicate information, requiring both cognitive and technical skills" (Digital Literacy Task Force, 2013, p. 2). Therefore, increasing EFL learners' digital literacy might further help the learners perform better in mobile-based vocabulary learning environments.

However, the findings of this study need to be generalised with caution as it was conducted among specific participants in EFL setting. Additionally, due to learner differences in cognition and English proficiency level, which could probably affect the way the learners in the present study were engaged in the mobile-based environments, the EFL researchers are recommended to consider more in-depth studies about EFL learners with similar cognition and English proficiency level. It is also worthwhile to propose that neither superiority nor priority of any special kind of mobile phone was taken into consideration during this study. And even the practicality and applicability of mobile phones over other mobile technologies, such as PDAs, iPads, or tablets, were not regarded as the vital emphasis of this study.

Furthermore, as we could not make sure if the participants used the applications consistently and if they did not involve in other modes of learning, other researchers might carry out a similar study and adopt an online programme and instrument to check and record EFL learners' use of the applications. This could be controlled somehow, but as the nature of the longitudinal study, other factors might be involved affecting learners' improvement in a long period of time which cannot be controlled completely. Moreover, as we only used one group of learners, other researchers may add a control group of learners who do not use mobile based applications to make a comparison between the experimental and control groups' vocabulary learning attitudes and self-regulatory capacity in vocabulary learning so that the findings could be generalised and more valid.

In the Iranian context, on the other hand, vocabulary is usually decontextualised and taught as a separate skill. As a result, EFL learners do not usually find opportunities to use the words in their speaking and writing activities. That is, following Li and Hafner (2022), instructors in the EFL contexts only try to enhance EFL learners' receptive knowledge of vocabulary by teaching vocabularies individually or through listening and reading skills; however, instructors ignore EFL learners' productive knowledge of vocabulary. This means that in the EFL contexts, learners are not usually involved in speaking and writing activities to use their vocabulary knowledge productively. As the current study indicated, mobile-based applications provide a user-friendly and convenient online environment for EFL student not only to improve their receptive knowledge of vocabulary, but also to apply their newly acquire vocabularies productively through collaborative activities with their peers.

## DATA AVAILABILITY STATEMENT

The data analyzed in this study is subject to the following licenses/restrictions: the dataset will be available upon request by contacting the corresponding author. Requests to access these datasets should be directed to jfathi13@yahoo.com.

## REFERENCES

- Azevedo, R., and Cromley, J. G. (2004). Does training on self-regulated learning facilitate students' learning with hypermedia? *J. Educ. Psychol.* 96, 523–535. doi: 10.1037/0022-0663.96.3.523
- Barkaoui, K. (2014). Quantitative approaches for analysing longitudinal data in second language research. *Annu. Rev. Appl. Linguist.* 34, 65–101. doi: 10.1017/S0267190514000105
- Barth, I., Zou, B., Spector-Cohen, E., and Sitman, R. (2020). "Supporting self-directed vocabulary study: towards effective design of OERs for mobile learning environments," in *Recent Developments in Technology-Enhanced and Computer-Assisted Language Learning* (IGI Global), 48–71.
- Başoğlu, E., and Akdemir, O. (2010). A comparison of undergraduate students' English vocabulary learning: using mobile phones and flash cards. *Turk. Online J. Educ. Technol.* 9, 1–7.
- Burston, J. (2014). MALL: the pedagogical challenges. *Comput. Assist. Lang. Learn.* 27, 344–357. doi: 10.1080/09588221.2014.914539
- Burston, J. (2015). Twenty years of MALL project implementation: a meta-analysis of learning outcomes. *ReCALL* 27, 4–20. doi: 10.1017/S0958344014000159
- Çakmak, F., and Erçetin, G. (2018). Effects of gloss type on text recall and incidental vocabulary learning in mobile-assisted L2 listening. *ReCALL* 30, 24–47. doi: 10.1017/S0958344017000155
- Carneiro, R., Lefrere, P., and Steffens, K. (eds.) (2007). *Self-Regulated Learning in Technology Enhanced Learning Environments: A European Review*. London: Kaleidoscope Network of Excellence.
- Caruana, E. J., Roman, M., Hernández-Sánchez, J., and Solli, P. (2015). Longitudinal studies. *J. Thorac. Dis.* 7, 537–540.
- Çetinkaya, L., and Sütçü, S. S. (2018). The effects of Facebook and WhatsApp on success in English vocabulary instruction. *J. Comput. Assist. Learn.* 34, 504–514. doi: 10.1111/jcal.12255
- Chacón-Beltrán, R. (2018). Vocabulary learning strategies outside the classroom context: what adults learn in a technology-based learner-centred environment. *Lang. Learn. J.* 46, 583–593. doi: 10.1080/09571736.2018.1503135
- Chang, C. K., and Hsu, C. K. (2011). A mobile-assisted synchronously collaborative translation annotation system for English as a foreign language (EFL) reading

## ETHICS STATEMENT

The studies involving human participants were reviewed and approved by University of Kurdistan. The patients/participants provided their written informed consent to participate in this study.

## AUTHOR CONTRIBUTIONS

All authors have contributed equally to data collection, data analysis, research questions, topic development, writing the manuscript as well as its revision, and language editing.

## FUNDING

This paper was supported by Blended Learning Design and Practice of EFL Courses based on BOPPPS and an APP Named Chaoxing Learning (ZH202138), A Research on the Teaching Effect Evaluation of Ideological and Political Education in Medical Humanities (2021RWJD01SZ), and A Research on the Teaching Design and Application of Undergraduate Dissertations in English Majors (202102431018).

- comprehension. *Comput. Assist. Lang. Learn.* 24, 155–180. doi: 10.1080/09588221.2010.536952
- Chen, C. M., Chen, L. C., and Yang, S. M. (2019). An English vocabulary learning app with self-regulated learning mechanism to improve learning performance and motivation. *Comput. Assist. Lang. Learn.* 32, 237–260. doi: 10.1080/09588221.2018.1485708
- Chen, Y. L., and Hsu, C. C. (2020). Self-regulated mobile game-based English learning in a virtual reality environment. *Comput. Educ.* 154, 103910–103915. doi: 10.1016/j.compedu.2020.103910
- Chen, W., Tan, N. Y. L., Looi, C.-K., Zhang, B., and Seow, P. S. K. (2008). Handheld computers as cognitive tools: technology-enhanced environmental learning. *Res. Pract. Technol. Enhanc. Learn.* 3, 231–252. doi: 10.1142/S1793206808000513
- Chen, M. P., Wang, L. C., Zou, D., Lin, S. Y., Xie, H., and Tsai, C. C. (2020). Effects of captions and English proficiency on learning effectiveness, motivation and attitude in augmented-reality-enhanced theme-based contextualised EFL learning. *Comput. Assist. Lang. Learn.* 35, 381–411. doi: 10.1080/09588221.2019.1704787
- Cheng, C. H., and Chen, C. H. (2022). Investigating the impacts of using a mobile interactive English learning system on the learning achievements and learning perceptions of student with different backgrounds. *Comput. Assist. Lang. Learn.* 35, 88–113. doi: 10.1080/09588221.2019.1671460
- Cheung, W. S., and Hew, K. F. (2009). A review of research methodologies used in studies on mobile handheld devices in K-12 and higher education settings. *Australas. J. Educ. Technol.* 25, 153–183. doi: 10.14742/ajet.1148
- Chinnery, G. M. (2006). Going to the MALL: Mobile assisted language learning. *Lang. Learn. Technol.* 10, 9–16.
- Chu, H. C., Wang, C. C., and Wang, L. (2019). Impacts of concept map-based collaborative Mobile gaming on English grammar learning performance and behaviours. *J. Educ. Technol. Soc.* 22, 86–100.
- Corno, L., and Mandinach, E. B. (1983). The role of cognitive engagement in classroom learning and motivation. *Educ. Psychol.* 18, 88–108. doi: 10.1080/00461528309529266
- Crompton, H. (2013). "A historical overview of mobile learning: towards learner-centered education," in *Handbook of Mobile Learning*, eds. Z. L. Berge and L. Y. Muilenburg, 3–14.
- Digital Literacy Task Force (2013). Digital Literacy, Libraries, and Public Policy. Report of the Office of Information Technology Policy's Digital Literacy



- Task Force. Available at: [http://www.districtdispatch.org/wp-content/uploads/2013/01/2012\\_OITP-digilitreport\\_1\\_22\\_13.pdf](http://www.districtdispatch.org/wp-content/uploads/2013/01/2012_OITP-digilitreport_1_22_13.pdf)
- Dörnyei, Z. (2005). *The Psychology of the Language Learner: Individual Differences in Second Language Acquisition*. New Jersey: Lawrence Erlbaum Associates.
- Dörnyei, Z. (2007). *Research Methods in Applied Linguistics*. Oxford: Oxford University Press.
- Duman, G., Orhon, G., and Gedik, N. (2015). Research trends in mobile assisted language learning from 2000 to 2012. *ReCALL* 27, 197–216. doi: 10.1017/S0958344014000287
- Ebadi, S., and Rahimi, M. (2017). Exploring the impact of online peer-editing using Google docs on EFL learners' academic writing skills: a mixed-methods study. *Comput. Assist. Lang. Learn.* 30, 787–815. doi: 10.1080/09588221.2017.1363056
- Ebadi, S., and Rahimi, M. (2018). An exploration into the impact of WebQuest-based classroom on EFL learners' critical thinking and academic writing skills: a mixed-methods study. *Comput. Assist. Lang. Learn.* 31, 617–651. doi: 10.1080/09588221.2018.1449757
- Ebadi, S., and Rahimi, M. (2019). Mediating EFL learners' academic writing skills in online dynamic assessment using Google docs. *Comput. Assist. Lang. Learn.* 32, 527–555. doi: 10.1080/09588221.2018.1527362
- Ellis, R. (2008). *The Study of Second Language Acquisition*. Oxford: Oxford University Press.
- Ellis, R. (2015). *Understanding Second Language Acquisition. 2nd Edn.* Oxford University Press.
- Engeström, Y. (1987). *Learning by Expanding: An Activity-Theoretical Approach to Developmental Research*. Helsinki: Orienta-Konsultit.
- Engeström, Y. (2001). Expansive learning at work: towards an activity-theoretical reconceptualization. *J. Educ. Work.* 14, 133–156. doi: 10.1080/13639080020028747
- Eubanks, J. F., Yeh, H. T., and Tseng, H. (2018). Learning Chinese through a twenty-first century writing workshop with the integration of mobile technology in a language immersion elementary school. *Comput. Assist. Lang. Learn.* 31, 346–366. doi: 10.1080/09588221.2017.1399911
- Fathi, J., Alipour, F., and Saeedian, A. (2018). Enhancing vocabulary learning and self-regulation via a mobile application: an investigation of the memrise app. *JMRELS* 5, 27–46.
- Fathi, J., Arabani, A. S., and Mohamadi, P. (2021). The effect of collaborative writing using Google docs on EFL learners' writing performance and writing self-regulation. *Lang. Relat. Res.* 12, 333–359.
- Fathi, J., and Rahimi, M. (2020). Examining the impact of flipped classroom on writing complexity, accuracy, and fluency: a case of EFL students. *Comput. Assist. Lang. Learn.* 1–39. doi: 10.1080/09588221.2020.1825097 (in press).
- Ferrer, E., and McArdle, J. J. (2010). Longitudinal modelling of developmental changes in psychological research. *Curr. Dir. Psychol. Sci.* 19, 149–154. doi: 10.1177/0963721410370300
- García Botero, G., Botero Restrepo, M. A., Zhu, C., and Questier, F. (2021). Complementing in-class language learning with voluntary out-of-class MALL. Does training in self-regulation and scaffolding make a difference? *Comput. Assist. Lang. Learn.* 34, 1013–1039. doi: 10.1080/09588221.2019.1650780
- Godwin-Jones, R. (2011). Mobile apps for language learning. *Lang. Learn. Technol.* 15, 2–11.
- Haq, G. R. E. (2019). A Study of Mobile-Assisted Language Learning for Self-Regulated Learning of Indonesian EFL Students. [PhD Thesis]. National Taiwan University of Science and Technology.
- Hernández, R., and Rankin, P. (2015). *Higher Education and Second Language Learning: Promoting Self-Directed Learning in New Technological and Educational Contexts*. Peter Lang.
- Hu, L. T., and Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: conventional criteria versus new alternatives. *Struct. Equ. Model. Multidiscip. J.* 6, 1–55. doi: 10.1080/10705519909540118
- Huang, C. S., Yang, S. J., Chiang, T. H., and Su, A. (2016). Effects of situated mobile learning approach on learning motivation and performance of EFL students. *J. Educ. Technol. Soc.* 19, 263–276.
- Hulstijn, J. H., and Laufer, B. (2001). Some empirical evidence for the involvement load hypothesis in vocabulary acquisition. *Lang. Learn.* 51, 539–558. doi: 10.1111/0023-8333.00164
- Hung, P. H., Hwang, G. J., Su, I. S., and Lin, I. H. (2012). A concept-map integrated dynamic assessment system for improving ecology observation competences in mobile learning activities. *Turk. Online J. Educ. Technol.* 11, 10–19.
- Hwang, G. J., and Fu, Q. K. (2018). Trends in the research design and application of mobile language learning: a review of 2007–2016 publications in selected SSCI journals. *Interact. Learn. Environ.* 27, 567–581. doi: 10.1080/10494820.2018.1486861
- Jia, C., and Hew, K. F. T. (2019). Supporting lower-level processes in EFL listening: the effect on learners' listening proficiency of a dictation program supported by a mobile instant messaging app. *Comput. Assist. Lang. Learn.* 35, 141–168. doi: 10.1080/09588221.2019.1671462
- Kennedy, C., and Levy, M. (2008). L'italiano al telefonino: using SMS to support beginners' language learning. *ReCALL* 20, 315–330. doi: 10.1017/S0958344008000530
- Kiernan, P. J., and Aizawa, K. (2004). Cell phones in task based learning: are cell phones useful language learning tools? *ReCALL* 16, 71–84. doi: 10.1017/S0958344004000618
- Kondo, M., Ishikawa, Y., Smith, C., Sakamoto, K., Shimomura, H., and Wada, N. (2012). Mobile assisted language learning in university EFL courses in Japan: developing attitudes and skills for self-regulated learning. *ReCALL* 24, 169–187. doi: 10.1017/S0958344012000055
- Kukulka-Hulme, A. (2006). "Mobile language learning now and in the future," in *Från en vision till praktik: Språkutbildning och informationsteknik (From Vision to Practice: Language Learning and IT)*. ed. P. Svensson (Swedish Net University (Nätuniversitetet)), 295–310.
- Kukulka-Hulme, A., and Traxler, J. (2005). *Mobile Learning: A Handbook for Educators and Trainers*. Psychology Press.
- Kurt, M., and Bensen, H. (2017). Six seconds to visualize the word: improving EFL learners' vocabulary through VVVs. *J. Comput. Assist. Learn.* 33, 334–346. doi: 10.1111/jcal.12182
- Lai, A. (2016). Mobile immersion: an experiment using mobile instant messenger to support second-language learning. *Interact. Learn. Environ.* 24, 277–290. doi: 10.1080/10494820.2015.1113706
- Lai, C., and Gu, M. (2011). Self-regulated out-of-class language learning with technology. *Comput. Assist. Lang. Learn.* 24, 317–335. doi: 10.1080/09588221.2011.568417
- Laurillard, D. (2007). "Pedagogical forms of mobile learning: framing research questions," in *Mobile Learning: Towards a Research Agenda*. ed. N. Pachler (WLE Centre, Institute of Education, University of London), 153–175.
- Levy, M., and Kennedy, C. (2005). "Learning Italian via mobile SMS," in *Mobile Learning: A Handbook for Educators and Trainers*. eds. A. Kukulka-Hulme and J. Traxler (Taylor & Francis), 76–83.
- Levy, M., and Stockwell, M. (2006). Effective use of CALL technologies: Finding the right balance. *Chang. Lang. Educ. CALL* 1, 301–320.
- Li, Y., and Hafner, C. A. (2022). Mobile-assisted vocabulary learning: investigating receptive and productive vocabulary knowledge of Chinese EFL learners. *ReCALL* 34, 66–80. doi: 10.1017/S0958344021000161
- Liaw, S. S., Hatala, M., and Huang, H. M. (2010). Investigating acceptance towards mobile learning to assist individual knowledge management: based on activity theory approach. *Comput. Educ.* 54, 446–454. doi: 10.1016/j.compedu.2009.08.029
- Lin, J. J., and Lin, H. (2019). Mobile-assisted ESL/EFL vocabulary learning: a systematic review and meta-analysis. *Comput. Assist. Lang. Learn.* 32, 878–919. doi: 10.1080/09588221.2018.1541359
- Lin, C. C., and Yu, Y. C. (2017). Effects of presentation modes on mobile-assisted vocabulary learning and cognitive load. *Interact. Learn. Environ.* 25, 528–542. doi: 10.1080/10494820.2016.1155160
- Liu, P. L. (2016). Mobile English vocabulary learning based on concept-mapping strategy. *Lang. Learn. Technol.* 20, 128–141.
- Liu, G. Z., Chen, J. Y., and Hwang, G. J. (2018). Mobile-based collaborative learning in the fitness center: a case study on the development of English listening comprehension with a context-aware application. *Br. J. Educ. Technol.* 49, 305–320. doi: 10.1111/bjet.12581
- Loewen, S., Crowther, D., Isbell, D. R., Kim, K. M., Maloney, J., Miller, Z. F., et al. (2019). Mobile-assisted language learning: a Duolingo case study. *ReCALL* 31, 293–311. doi: 10.1017/S0958344019000065
- Looi, C. K., Seow, P., Zhang, B., So, H. J., Chen, W., and Wong, L. H. (2010). Leveraging mobile technology for sustainable seamless learning: a research agenda. *Br. J. Educ. Technol.* 41, 154–169. doi: 10.1111/j.1467-8535.2008.00912.x

- Lu, M. (2008). Effectiveness of vocabulary learning via mobile phone. *J. Comput. Assist. Learn.* 24, 515–525. doi: 10.1111/j.1365-2729.2008.00289.x
- Masgoret, A. M., and Gardner, R. C. (2003). Attitudes, motivation, and second language learning: a meta-analysis of studies conducted by Gardner and associates. *Lang. Learn.* 53, 123–163. doi: 10.1111/1467-9922.00212
- McArdle, J. J. (2001). “A latent difference score approach to longitudinal dynamic structural analyses,” in *Structural Equation Modelling: Present and Future*. eds. R. Cudeck, S. du Toit and D. Sörbom (Lincolnwood, IL: Scientific Software International), 342–380.
- McArdle, J. J., and Hamagami, F. (2001). “Latent difference score structural models for linear dynamic analyses with incomplete longitudinal data,” in *New Methods for the Analysis of Change*. eds. L. M. Collins and A. G. Sayer (Washington, DC: APA Press), 139–175.
- McArdle, J. J., and Prindle, J. J. (2008). A latent change score analysis of a randomised clinical trial in reasoning training. *Psychol. Aging* 23, 702–719. doi: 10.1037/a0014349
- Melhuish, K., and Fallon, G. (2010). Looking to the future: M-learning with the iPad. *Computers in New Zealand Schools: Learning, Leading, Technology* 22, 1–16.
- Muthén, L. K., and Muthén, B. O. (2012). *Mplus User's Guide* (6th ed.). CA: Muthén & Muthén.
- Nation, I. S. P. (2001). *Learning Vocabulary in Another Language*. Cambridge University Press.
- Nation, I. S. P. (2020). “The different aspects of vocabulary knowledge,” in *The Routledge Handbook of Vocabulary Studies*. ed. S. Webb (Routledge), 15–29.
- Nation, I. S. P., and Webb, S. (2011). *Researching and Analysing Vocabulary*. Heinle.
- Ning, H. K., and Downing, K. (2010). The reciprocal relationship between motivation and self-regulation: a longitudinal study on academic performance. *Learn. Individ. Differ.* 20, 682–686. doi: 10.1016/j.lindif.2010.09.010
- Norris, C., Hossain, A., and Soloway, E. (2011). Using smartphones as essential tools for learning: a call to place schools on the right side of the 21st century. *Educ. Technol.* 51, 18–25.
- Oblinger, D., and Oblinger, J. (2005). *Educating the Net Generation*. Boulder: Colorado.
- Ortega, L., and Byrnes, H. (2009). *The Longitudinal Study of Advanced L2 Capacities: An Introduction*. Routledge, 19–36.
- Rachels, J. R., and Rockinson-Szapkiw, A. J. (2018). The effects of a mobile gamification app on elementary students' Spanish achievement and self-efficacy. *Comput. Assist. Lang. Learn.* 31, 72–89. doi: 10.1080/09588221.2017.1382536
- Rahimi, M., and Fathi, J. (2021). Exploring the impact of wiki-mediated collaborative writing on EFL students' writing performance, writing self-regulation, and writing self-efficacy: a mixed methods study. *Comput. Assist. Lang. Learn.* 1–48. doi: 10.1080/09588221.2021.1888753 (in press).
- Rassaei, E. (2019). Computer-mediated text-based and audio-based corrective feedback, perceptual style and L2 development. *System* 82, 97–110. doi: 10.1016/j.system.2019.03.004
- Rassaei, E. (2020). Effects of mobile-mediated dynamic and nondynamic glosses on L2 vocabulary learning: a sociocultural perspective. *Mod. Lang. J.* 104, 284–303. doi: 10.1111/modl.12629
- Rassaei, E. (2021). Implementing mobile-mediated dynamic assessment for teaching request forms to EFL learners. *Comput. Assist. Lang. Learn.* 1–31. doi: 10.1080/09588221.2021.1912105 (in press).
- Rosell-Aguilar, F. (2018). Autonomous language learning through a mobile application: a user evaluation of the busuu app. *Comput. Assist. Lang. Learn.* 31, 854–881. doi: 10.1080/09588221.2018.1456465
- Schmitt, N. (2010). *Researching Vocabulary: A Vocabulary Research Manual*. Palgrave Macmillan.
- Schmitt, N., and Schmitt, D. (2020). *Vocabulary in Language Teaching*. Cambridge University Press.
- Seibert Hanson, A. E., and Brown, C. M. (2020). Enhancing L2 learning through a mobile assisted spaced-repetition tool: an effective but bitter pill? *Comput. Assist. Lang. Learn.* 33, 133–155. doi: 10.1080/09588221.2018.1552975
- Sha, L., Looi, C. K., Chen, W., Seow, P., and Wong, L. H. (2012). Recognizing and measuring self-regulated learning in a mobile learning environment. *Comput. Hum. Behav.* 28, 718–728. doi: 10.1016/j.chb.2011.11.019
- Shadiev, R., Hwang, W. Y., and Huang, Y. M. (2017). Review of research on mobile language learning in authentic environments. *Comput. Assist. Lang. Learn.* 30, 284–303. doi: 10.1080/09588221.2017.1308383
- Sharples, M. (2006). *Big Issues in Mobile Learning*. In *Report of a Workshop by the Kaleidoscope Network of Excellence Mobile Learning Initiative*. LSRI, University of Nottingham.
- Shea, P., and Bidjerano, T. (2010). Learning presence: towards a theory of self-efficacy, self-regulation, and the development of a communities of inquiry in online and blended learning environments. *Comput. Educ.* 55, 1721–1731. doi: 10.1016/j.compedu.2010.07.017
- Simpson, J. (2005). Learning electronic literacy skills in an online language learning community. *Comput. Assist. Lang. Learn.* 18, 327–345. doi: 10.1080/09588220500335463
- Song, Y., and Fox, R. (2008). Using PDA for undergraduate student incidental vocabulary testing. *ReCALL* 20, 290–314. doi: 10.1017/S0958344008000438
- Soyoof, A., Reynolds, B. L., Shadiev, R., and Vazquez-Calvo, B. (2022). A mixed-methods study of the incidental acquisition of foreign language vocabulary and healthcare knowledge through serious game play. *Comput. Assist. Lang. Learn.* 1–34. doi: 10.1080/09588221.2021.2021242 (in press).
- Soyoof, A., Reynolds, B. L., Vazquez-Calvo, B., and McLay, K. (2021). Informal digital learning of English (IDLE): a scoping review of what has been done and a look towards what is to come. *Comput. Assist. Lang. Learn.* 1–27. doi: 10.1080/09588221.2021.1936562 (in press).
- Stockwell, G. (2010). Using mobile phones for vocabulary activities: examining the effect of platform. *Lang. Learn. Technol.* 14, 95–110.
- Sun, Z., Lin, C. H., You, J., Shen, H. J., Qi, S., and Luo, L. (2017). Improving the English-speaking skills of young learners through mobile social networking. *Comput. Assist. Lang. Learn.* 30, 304–324. doi: 10.1080/09588221.2017.1308384
- Sung, Y. T., Chang, K. E., and Liu, T. C. (2016). The effects of integrating mobile devices with teaching and learning on students' learning performance: a meta-analysis and research synthesis. *Comput. Educ.* 94, 252–275. doi: 10.1016/j.compedu.2015.11.008
- Sung, Y. T., Chang, K. E., and Yang, J. M. (2015). How effective are mobile devices for language learning? A meta-analysis. *Educ. Res. Rev.* 16, 68–84. doi: 10.1016/j.edurev.2015.09.001
- Teng, F. (2020). Vocabulary learning through videos: captions, advance-organizer strategy, and their combination. *Comput. Assist. Lang. Learn.* 35, 518–550. doi: 10.1080/09588221.2020.1720253
- Thornton, P., and Houser, C. (2005). Using mobile phones in English education in Japan. *J. Comput. Assist. Learn.* 21, 217–228. doi: 10.1111/j.1365-2729.2005.00129.x
- Tseng, W. T., Dörnyei, Z., and Schmitt, N. (2006). A new approach to assessing strategic learning: the case of self-regulation in vocabulary acquisition. *Appl. Linguis.* 27, 78–102. doi: 10.1093/applin/ami046
- Tseng, W. T., and Schmitt, N. (2008). Towards a model of motivated vocabulary learning: a structural equation modelling approach. *Lang. Learn.* 58, 357–400. doi: 10.1111/j.1467-9922.2008.00444.x
- Uden, L. (2007). Activity theory for designing mobile learning. *Int. J. Mob. Learn. Organ.* 1, 81–102. doi: 10.1504/IJMLLO.2007.011190
- Vygotsky, L. (1978). *Mind in Society: The Development of Higher Mental Processes*. Cambridge, MA: Harvard University Press.
- Wali, E., Winters, N., and Oliver, M. (2008). Maintaining, changing and crossing contexts: an activity theoretic reinterpretation of mobile learning. *Res. Learn. Technol.* 16, 41–57. doi: 10.3402/rlt.v16i1.10884
- Wang, H. Y., Liu, G. Z., and Hwang, G. J. (2017). Integrating socio-cultural contexts and location-based systems for ubiquitous language learning in museums: a state of the art review of 2009–2014. *Br. J. Educ. Technol.* 48, 653–671. doi: 10.1111/bjet.12424
- Wang, Y. H., and Shih, S. K. H. (2015). Mobile-assisted language learning: effects on EFL vocabulary learning. *Int. J. Mob. Commun.* 13, 358–375. doi: 10.1504/IJMC.2015.070060
- Wong, L. H. (2012). A learner-centric view of mobile seamless learning. *Br. J. Educ. Technol.* 43, E19–E23. doi: 10.1111/j.1467-8535.2011.01245.x
- Wong, L. H., and Looi, C. K. (2010). Vocabulary learning by mobile-assisted authentic content creation and social meaning-making: two case studies. *J. Comput. Assist. Learn.* 26, 421–433. doi: 10.1111/j.1365-2729.2010.00357.x
- Wongsuriya, P. (2020). Improving the Thai students' ability in English pronunciation through mobile application. *Educ. Res. Rev.* 15, 175–185. doi: 10.5897/ERR2020.3904

- Wu, Q. (2015). Designing a smartphone app to teach English (L2) vocabulary. *Comput. Educ.* 85, 170–179. doi: 10.1016/j.compedu.2015.02.013
- Wyra, M., and Lawson, M. J. (2018). Foreign language vocabulary learning using the keyword method: strategy and meta-strategy knowledge. *Lang. Learn. J.* 46, 605–621. doi: 10.1080/09571736.2018.1503138
- Xu, Q., and Peng, H. (2017). Investigating mobile-assisted oral feedback in teaching Chinese as a second language. *Comput. Assist. Lang. Learn.* 30, 173–182. doi: 10.1080/09588221.2017.1297836
- Zhang, H., Song, W., and Burston, J. (2011). Re-examining the effectiveness of vocabulary learning via mobile phones. *Turk. Online J. Educ. Technol.* 10, 203–214.
- Zhang, R., and Zou, D. (2020). Types, purposes, and effectiveness of state-of-the-art technologies for second and foreign language learning. *Comput. Assist. Lang. Learn.* 1–47. doi: 10.1080/09588221.2020.1744666 (in press).
- Zheng, C., Liang, J. C., Li, M., and Tsai, C. C. (2018). The relationship between English language learners' motivation and online self-regulation: a structural equation modelling approach. *System* 76, 144–157. doi: 10.1016/j.system.2018.05.003
- Zimmerman, B. J. (2000). "Attainment of self-regulation: a social cognitive perspective," in *Handbook of Self-Regulation*. eds. M. Boekaerts, P. R. Pintrich and M. Zeidner (Academic Press), 13–41.
- Zimmerman, B. J. (2008). "Goal setting: a key proactive source of academic self-regulation," in *Motivation and Self-Regulated Learning: Theory, Research and Applications*. eds. D. H. Schunk and B. J. Zimmerman (Lawrence Erlbaum), 267–295.
- Zimmerman, B. J., and Schunk, D. H. (2008). "Motivation: an essential dimension of self-regulated learning," in *Motivation and Self-Regulated Learning: Theory, Research, and Applications*. eds. D. H. Schunk and B. J. Zimmerman (Lawrence Erlbaum Associates), 1–30.
- Zou, D., Huang, Y., and Xie, H. (2021). Digital game-based vocabulary learning: where are we and where are we going? *Comput. Assist. Lang. Learn.* 34, 751–777. doi: 10.1080/09588221.2019.1640745
- Conflict of Interest:** The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.
- Publisher's Note:** All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Copyright © 2022 Lei, Fathi, Noorbakhsh and Rahimi. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.



# Using Mobile Devices for Vocabulary Learning Outside the Classroom: Improving the English as Foreign Language Learners' Knowledge of High-Frequency Words

Azadeh Rahmani<sup>1</sup>, Vahid Asadi<sup>2</sup> and Ismail Xodabande<sup>3\*</sup>

<sup>1</sup> Department of Foreign Languages, Danesh Alborz University, Takestan, Iran, <sup>2</sup> Department of Foreign Languages, Islamic Azad University, Yasuj, Iran, <sup>3</sup> Department of Foreign Languages, Kharazmi University, Tehran, Iran

## OPEN ACCESS

### Edited by:

Xuesong Gao,  
University of New South Wales,  
Australia

### Reviewed by:

Tao Hao,  
East China Normal University, China  
Ghazwan Mohammed,  
University of Bisha, Saudi Arabia

### \*Correspondence:

Ismail Xodabande  
ismail.kh.tefl@gmail.com

### Specialty section:

This article was submitted to  
Language Sciences,  
a section of the journal  
Frontiers in Psychology

Received: 19 March 2022

Accepted: 07 June 2022

Published: 24 June 2022

### Citation:

Rahmani A, Asadi V and  
Xodabande I (2022) Using Mobile  
Devices for Vocabulary Learning  
Outside the Classroom: Improving  
the English as Foreign Language  
Learners' Knowledge  
of High-Frequency Words.  
Front. Psychol. 13:899885.  
doi: 10.3389/fpsyg.2022.899885

The present study investigated the impacts of mobile assisted vocabulary learning via digital flashcards (DFs). The data were collected from 44 adult English as Foreign Language (EFL) learners in three intact classes in a private language teaching institute in Iran, randomly assigned to experimental ( $N = 27$ ) and control ( $N = 17$ ) learning conditions. The experimental group used a freely available DF application (i.e., NGSL builder) to learn items from a recently developed corpus-based word list for high-frequency vocabulary in English (NGSL). The treatment was implemented as out-of-the-classroom learning activities where the EFL learners used DFs to augment their vocabulary knowledge, and their learning gains were compared to the control group that received regular English language education. The participants' vocabulary knowledge was tested in pre-, post-, and delayed post-tests, and the findings indicated that using DFs for outside the classroom vocabulary learning contributed significantly to short- and long-term improvements in the knowledge of high-frequency words. The study provided empirical evidence for the affordances of mobile assisted vocabulary learning for learning a considerable proportion of core vocabulary and has some implications for addressing the vocabulary learning needs of EFL learners.

**Keywords:** mobile assisted language learning, digital flashcards, vocabulary – general vocabulary, English as a foreign language, word list learning

## INTRODUCTION

The importance of vocabulary in second language literacy developments is widely acknowledged (Nation, 2013; Webb and Nation, 2017). More specifically, vocabulary knowledge is considered crucial for language learners' proficiency development in the four macro skills including reading, listening, speaking, and writing (Clenton and Booth(eds), 2020), and regarded to be as the single most important variable in language learning in general (Dodigovic and Agustín-Llach, 2020). In light of these considerations, a major pedagogical challenge in teaching English to speakers of other languages (TESOL) has been addressing the vocabulary learning needs of learners at different stages of language learning (Nation, 2013). Over the past decades, this consistent concern



resulted in the creation of a number of corpus-based word lists for high-frequency and academic vocabulary to inform language teaching materials developments (Coxhead, 2000; Gardner and Davies, 2014; Brezina and Gablasova, 2015; Browne, 2021; Cobb and Laufer, 2021). High-frequency or general service words are commonly used vocabulary items in spoken and written language that provide from 80 to 90% coverage in most texts in English. The academic vocabulary, however, refers to a group of words for describing abstract ideas and processes in academic discourse (Coxhead, 2000). Given the large coverage and hence the importance of high-frequency words, there is a need for systematically focusing on teaching them and finding effective strategies to facilitate their acquisition by language learners. Nonetheless, vocabulary instruction generally remained underestimated in second language classrooms (Dodigovic and Agustín-Llach, 2020), and language teachers mostly pay insufficient attention to it for some reasons such as shortage of classroom time and high workloads (Webb and Nation, 2017).

With significant developments in Information and Communication Technologies, there is a growing interest in TESOL to use various affordances provided by digital technologies for vocabulary learning (Burston and Giannakou, 2021; Hao et al., 2021; Yang et al., 2021; Yu and Trainin, 2021). The expanding body of knowledge in this line of inquiry indicated that technology assisted learning is generally facilitative for second language vocabulary development (Mahdi, 2017; Lin and Lin, 2019). Accordingly, the integration of mobile assisted vocabulary learning in TESOL provide opportunities for a systematic focus on teaching the vocabulary items that are essential for the communication needs of language learners (such as high-frequency words; Xodabande et al., 2022). Additionally, the implementation of technological resources to augment vocabulary development might be more effective for outside the classroom learning conditions, as in most cases formal educational settings are resistant to integrating new technologies into the established curricula and teaching methodologies (Lai and Gu, 2011). Furthermore, a unique feature of digital technologies with considerable potential for transforming education is the possibility of extending learning and teaching to anytime and anyplace. In this regard, complementing classroom instruction with beyond the classroom technology assisted activities gives more freedom to students to conduct self-directed learning and also helps them to become more autonomous learners which is a valuable educational asset (Xodabande et al., 2022).

Among the various strategies developed for learning words, flashcards provide learners with a fast and effective way to improve their vocabulary knowledge (Nation, 2013). With the growing availability of smartphone devices for language learners, digital flashcards (DFs) attracted increased interest from researchers in recent years (Sage et al., 2019; Xodabande and Atai, 2020; Yüksel et al., 2020; Xodabande et al., 2022). Collectively, research in this area indicated that DFs used on mobile or computer devices contributed significantly to language learners' vocabulary development and increased their motivation and engagement with the learning materials. DFs also proved to be more effective than paper- or computer-based cards mainly

because of their accessibility and portability as learning tools. Researchers also investigated the effectiveness of DFs for learning academic and technical vocabulary, and also as an instrument for autonomous and self-directed vocabulary learning. Despite significant positive learning outcomes, most of the studies in this line of research were conducted in short time periods with interventions targeting a limited number of vocabulary items (Xodabande and Atai, 2020). These limitations make it difficult to understand the long-term impacts of mobile assisted vocabulary learning in general (Lin and Lin, 2019), and the potential of such emerging learning environments for learning a large number of words in particular (Xodabande et al., 2022). Accordingly, there is a need for more empirical research to explore the affordances of mobile devices and DFs for learning vocabulary items in the corpus-based word lists, and to track the learning outcomes in the long term through delayed post-tests.

The present study aimed to facilitate the learning of high-frequency vocabulary by English as Foreign Language (EFL) learners *via* mobile devices and DFs. To this end, the high-frequency vocabulary in English is operationalized based on New General Service List (NGSL) project (Browne, 2021) which is among the most recently developed corpus-based core vocabulary lists. The NGSL contains 2,818 items (i.e., flemmas) derived from a 273 million-word subsection of Cambridge English Corpus. The pedagogical value of the NGSL arises from its considerable coverage (around 92.34%) of the corpus used for developing the list. Accordingly, mastery over the NGSL is essential for achieving the minimum threshold for text comprehension in English (Laufer, 2013), which might be regarded as the first step in developing the lexical knowledge of EFL learners. Furthermore, given the existing research-practice gap in corpus-based language education (Chambers, 2019), the current study also aimed to connect research in wordlist development to the language learning needs of EFL learners and investigated both short- and long-term impacts of mobile assisted vocabulary learning. This is particularly noteworthy as in most EFL learning contexts, many language learners fail to acquire the high-frequency vocabulary after years of language education (Webb and Chang, 2012), and interventions that have lasting impacts on their vocabulary knowledge contribute significantly to their second language literacy development. The following research questions were addressed in this study:

1. Does using DFs for outside the classroom mobile assisted vocabulary learning result in boosting EFL learners' vocabulary knowledge?
2. Does mobile assisted vocabulary learning have long-term impacts on vocabulary learning as assessed by a delayed post-test?

## METHOD

### Participants

Data for the present study were collected from three intact classes in a private language teaching institute in Iran. The participants were 44 adult language learners (16 males, 25 females) selected

based on their availability in the research context, and their ages ranged from 17 to 34 ( $M = 23$ ). According to the results of the Cambridge online placement test (Test Your English, 2022), most of the participants were in B1 level (pre-intermediate) proficiency in English as outlined in the Common European Framework of Reference (Council of Europe, 2001). Two classes were randomly assigned to the experimental conditions ( $N = 27$ ) and one class acted as a control group ( $N = 17$ ). In compliance with ethical considerations in educational research, informed consents were obtained from the participants and they were informed regarding the goals of the research, stages of data collection, and confidentiality of personal information collected in the study.

## Materials and Testing Instruments

The study used a freely available DFs application (i.e., NGSL builder; EFL Technologies, 2017) which is designed for learning high-frequency vocabulary items in the NGSL (Browne, 2021). The NGSL builder application allows learners to select vocabulary items in different levels from beginner (0% skipped) to advanced (75% skipped). The application also supports using DFs in receptive (meaning recall) and productive (form recall) learning modes. Additionally, the NGSL provides part of speech information and North American pronunciation for the words and has a built-in feature for recycling learned items in specific time intervals (i.e., spaced repetition). After an initial assessment of the vocabulary knowledge in the pre-test, 50% of NGSL words were skipped for the participants. For testing the vocabulary knowledge of the participants before and after the treatment, the study used NGSL Test (Stoeckel and Bennett, n.d.) which is a standard and validated measure of the NGSL with high reliability (Cronbach's  $\alpha = 0.90$ ). The test contains 100 items grouped in five levels and each level contains 20 multiple response questions.

## Procedures and Data Analysis

Prior to starting the treatment, the participants' vocabulary knowledge was tested using the NGSLT. Following this initial vocabulary learning needs identification, those participants in the experimental learning condition (classes A and B) installed the NGSL Builder application on their smartphone devices. The participants in these two classes received 2 h of training (in two online sessions) on using DFs for outside the classroom vocabulary learning and learned how to email their progress to the researchers through a feature available in the NGSL builder app. The treatment was implemented in a 3-month period where the participants were enrolled in a language learning course (focusing on developing communicative language ability) in a private language teaching institute in the context of the study. Accordingly, the participants in the control group (class C) received their regular education, while the participants in the experimental group used the NGSL builder application outside the classroom as a tool for boosting their vocabulary knowledge. The three classes had one and the same teacher, were taught based on a pre-specified syllabus determined by the language teaching institute, and identical materials and methods were used to teach the classes. The vocabulary knowledge of the participants was tested again at the end of the training period, and a follow-up delayed post-test was administered 3 months later. The scores on

vocabulary tests were analyzed using IBM SPSS (version 25) for descriptive and inferential statistics. In this regard, after running descriptive statistics to identify mean scores for different learning conditions, a mixed between-within subjects analysis of variance (ANOVA; Pallant, 2016) was performed to further investigate the main effects for each of the independent variables (i.e., group with two levels and time with three levels).

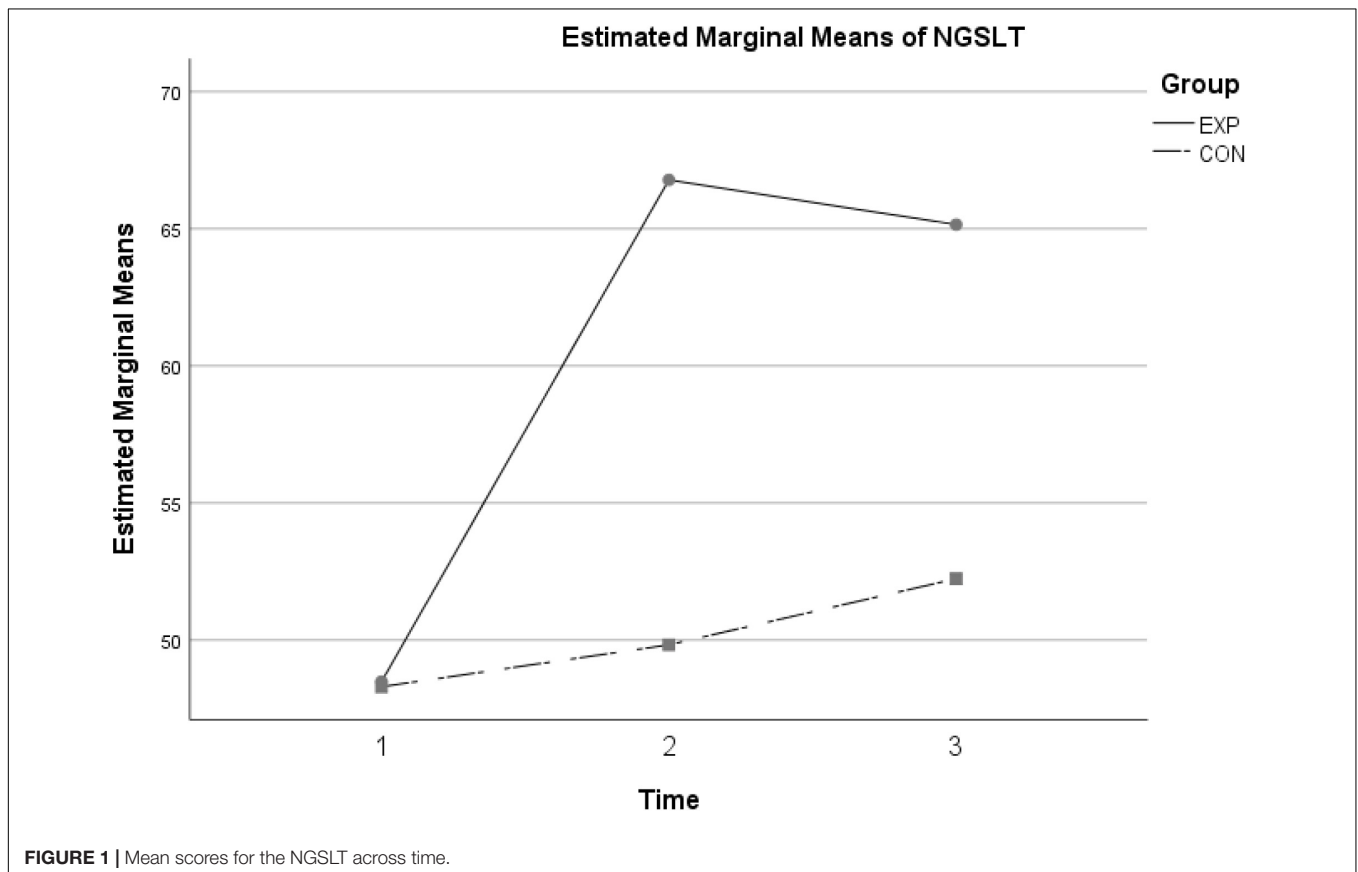
## RESULTS

The results of the descriptive statistics for the obtained scores by the two groups indicated that for test 1 (i.e., pre-test), the participants in the experimental and control group had similar scores on NGSLT, and the total mean value was calculated as 48.41 ( $SD = 5.16$ ). In order to check for any possible differences in vocabulary knowledge prior to treatment, an independent samples  $t$ -test was conducted that yielded non-significant results [ $t(42) = 0.116, p = 0.908$ , two tailed]. Nevertheless, in test 2 (i.e., post-test), the participants in the experimental group obtained higher scores ( $M = 66.78, SD = 4.31$ ) compared to the control group ( $M = 49.82, SD = 5.97$ ). Similar results were obtained in test 3 (i.e., delayed post-test) and the participants in the experimental group ( $M = 65.15, SD = 4.71$ ) scored higher than the control group ( $M = 52.24, SD = 6.33$ ). In order to further analyze the scores for statistical differences, a mixed between-within subjects ANOVA was conducted. Prior to investigating the main effects, preliminary analyses including Levene's test of equality of error variances (assumption of homogeneity of variances) and Box's test of equality of covariance matrices (assumption equality of covariance matrices) were examined and no violation of the assumptions required for ANOVA was observed. The results obtained for multivariate tests (**Table 1**) revealed a significant interaction effect between the two variables (i.e., time and group; Wilks'  $\Lambda = 0.39, F(2, 41) = 32.04, p = 0.000, \eta_p^2 = 0.61$ ). This interaction effect means that the changes in the obtained scores over time were different for the experimental and control groups (see **Figure 1** below). Moreover, the findings indicated that there is a statistically significant main effect for time (within subjects variable), Wilks'  $\Lambda = 0.24, F(2, 41) = 64.99, p \leq 0.001, \eta_p^2 = 0.76$ . As a result, the changes in the scores were statistically significant across the three testing times, and the findings revealed a very large effect size for the differences (Cohen, 1988). After investigating the main effect for time, the analysis proceeded to explore the between-subjects effects. As it is shown in **Table 2**, the main effect resulting from the two learning conditions was statistically significant,  $F(1, 42) = 117.56, p \leq 0.001, \eta_p^2 = 0.74$ . Accordingly, the findings indicated that mobile assisted vocabulary learning resulted in significant improvement in the vocabulary knowledge of the participants in the experimental group and the effect size of the difference was very large.

**Figure 1** provides visual representation of the scores obtained by the participants over time. Given the significant interaction effect for the pre-test to the post-test, this plot is helpful in interpreting the observed main effects. Accordingly, from the scores obtained by the experimental group is significantly

**TABLE 1** | Multivariate tests<sup>a</sup>.

Effect		Value	<i>F</i>	Hypothesis df	Error df	Sig.	Partial Eta Squared
Time	Pillai's Trace	0.760	64.991 <sup>b</sup>	2.000	41.000	0.000	0.760
	Wilks' Lambda	0.240	64.991 <sup>b</sup>	2.000	41.000	0.000	0.760
	Hotelling's Trace	3.170	64.991 <sup>b</sup>	2.000	41.000	0.000	0.760
	Roy's Largest Root	3.170	64.991 <sup>b</sup>	2.000	41.000	0.000	0.760
Time × Group	Pillai's Trace	0.610	32.044 <sup>b</sup>	2.000	41.000	0.000	0.610
	Wilks' Lambda	0.390	32.044 <sup>b</sup>	2.000	41.000	0.000	0.610
	Hotelling's Trace	1.563	32.044 <sup>b</sup>	2.000	41.000	0.000	0.610
	Roy's Largest Root	1.563	32.044 <sup>b</sup>	2.000	41.000	0.000	0.610

<sup>a</sup>Design: Intercept + Group. Within subjects design: Time.<sup>b</sup>Exact statistic.

larger. Moreover, from the post-test to the delayed post-test, the scores obtained by the experimental group dropped slightly, nonetheless, the associated mean value was still considerably higher than the mean score obtained by the control group.

In order to investigate the long-term impacts of the intervention, a series of pairwise comparisons were conducted for the scores obtained by the experimental and control groups on NGS LT (Table 3). The results indicated that for the experimental group, the scores on both post- and delayed post-tests were significantly higher than the scores obtained on the pre-test. However, the observed decline in the scores from post-test to the delayed post-test was not statistically significant. As for the control group, the findings indicated that the only significant

differences in the scores were between the pre-test and the delayed post-test. These findings provided empirical evidence for the short- and long-term effects of mobile assisted vocabulary learning *via* DFs.

## DISCUSSION AND CONCLUSION

The first research question investigated the contribution of mobile assisted vocabulary learning in augmenting the EFL learners' knowledge of high-frequency vocabulary in English. The findings revealed that outside the classroom vocabulary learning with DFs was effective in boosting the knowledge of core words

**TABLE 2 |** Tests of between-subjects effects.

Measure: NGSLT						
Transformed variable: Average						
Source	Type III sum of squares	df	Mean square	F	Sig.	Partial Eta Squared
Intercept	380,422.010	1	380,422.010	14,238.700	0.000	0.997
Group	3,140.919	1	3,140.919	117.561	0.000	0.737
Error	1,122.134	42	26.717			

**TABLE 3 |** Pairwise comparisons.

Measure: NGSLT							
Group	(I) TIME	(J) TIME	Mean difference (I-J)	Std. error	Sig. <sup>a</sup>	95% Confidence interval for difference <sup>a</sup>	
						Lower bound	Upper bound
EXP	1	2	-18.296*	1.584	0.000	-22.350	-14.243
		3	-16.667*	1.377	0.000	-20.192	-13.142
	2	1	18.296*	1.584	0.000	14.243	22.350
		3	1.630	1.369	0.734	-1.874	5.134
	3	1	16.667*	1.377	0.000	13.142	20.192
		2	-1.630	1.369	0.734	-5.134	1.874
CON	1	2	-1.529	1.760	1.000	-6.233	3.174
		3	-3.941*	1.048	0.005	-6.743	-1.139
	2	1	1.529	1.760	1.000	-3.174	6.233
		3	-2.412	2.216	0.878	-8.336	3.513
	3	1	3.941*	1.048	0.005	1.139	6.743
		2	2.412	2.216	0.878	-3.513	8.336

Based on estimated marginal means.

\*The mean difference is significant at the 0.05 level.

<sup>a</sup>Adjustment for multiple comparisons: Bonferroni.

in English, and the learning outcomes were considerable given the treatment time. More specifically, the participants in the experimental group learned around 18.3% of the NGS items over the course of 3 months which is equivalent to 500 words. The findings are congruent with previous studies that reported positive learning gains for DFs (Sage et al., 2019; Yüksel et al., 2020; Xodabande et al., 2022). There might be a number of factors contributing to such improvements in vocabulary knowledge. First, these learning gains mostly resulted from the retrieval practice which involves recalling the meaning of the target words from previous encounters (Barcroft, 2007). Such retrievals are associated with developments in both receptive and productive vocabulary knowledge in using paper and DFs (Li and Hafner, 2022). Second, language learners in the experimental group were exposed to a large number of words presented to them in the form of DFs, and given the effectiveness of explicit focus on vocabulary learning (Webb and Nation, 2017), they learned more words compared to the control group whose encounter with words was limited to the materials used in the classes. The use of mobile devices and DFs for vocabulary learning beyond the classroom might have also resulted in increased motivation among the participants in the experimental group, and this inherent motivation associated with new technologies (Stockwell, 2013) impacted learning outcomes.

The second research question was concerned with the long-term effects of mobile assisted vocabulary learning. The results of the delayed post-test indicated that the participants in the experimental group retained most of their acquired knowledge after 3 months following the treatment (albeit with some decline in their vocabulary test scores). This is also in line with the earlier studies that investigated the long-term impacts of DFs for vocabulary learning (Xodabande et al., 2022). However, the findings of the current study differ slightly from Xodabande and Atai (2020) who found a significant decline in the knowledge of academic vocabulary in the delayed post-test. One reason for this observation might be the nature of target vocabulary items, as the current study focused on high-frequency vocabulary which is more common than academic vocabulary in different text types. This frequency factor enhances the chance of more and varied encounters with target words for EFL learners which is not the case for academic vocabulary which is associated mostly with academic discourse. Additionally, the employed DFs application had a built-in feature for spaced repetition that further contributed to this long-term learning outcomes. Moreover, the EFL learners have more chance to use high-frequency vocabulary than the academic words, and meaningful use of words is essential for long-term improvements in vocabulary knowledge (Webb and Nation, 2017).



The study has some implications for integrating mobile assisted vocabulary learning into TESOL programs. First, as the findings indicated, complementing regular English courses with outside the classroom vocabulary learning with DFs facilitates the bridging of the vocabulary gap commonly reported for EFL learners (Webb and Chang, 2012). Accordingly, instructors might consider using such interventions to help their students in building a good vocabulary base despite limitations in classroom time and language teaching materials. Second, given the considerable learning outcomes (i.e., 18% of the NGSL) for learning with DFs, it seems that mobile assisted vocabulary learning brings opportunities for faster growth in the lexical dimension of second language literacy development which normally takes many years (Webb and Chang, 2012; Nation, 2013; Webb and Nation, 2017). This is specifically important in EFL contexts where learners need to master high-frequency vocabulary as a foundation for language learning in general (Clenton and Booth(eds), 2020; Dodigovic and Agustín-Llach, 2020). Furthermore, despite some limitations in sampling procedures and being concerned with only receptive knowledge of vocabulary, the current study is an example of the efforts made by language teachers to connect findings from corpus-based studies of language to the vocabulary learning needs of EFL learners. Such undertakings might be more effective if mobile assisted learning becomes more integrated into instructional curricula and teaching

methodologies. Further research might consider the impacts of using DFs on developing other aspects of vocabulary knowledge such as productive use of learned items or their collocations.

## DATA AVAILABILITY STATEMENT

The original contributions presented in the study are included in the article/supplementary material; further inquiries can be directed to the corresponding author.

## ETHICS STATEMENT

Ethical review and approval was not required for the study on human participants in accordance with the local legislation and institutional requirements. The patients/participants provided their written informed consent to participate in this study.

## AUTHOR CONTRIBUTIONS

All authors contributed to the design and implementation of the research, to the analysis of the results, and to the writing of the manuscript.

## REFERENCES

- Barcroft, J. (2007). Effects of opportunities for word retrieval during second language vocabulary learning. *Lang. Learn.* 57, 35–56. doi: 10.1111/j.1467-9922.2007.00398.x
- Brezina, V., and Gablasova, D. (2015). Is there a core general vocabulary? introducing the new general service list. *Appl. Ling.* 36, 1–22. doi: 10.1093/applin/amt018
- Browne, C. (2021). *The NGSL Project: Building Wordlists and Resources to help EFL Learners (and Teachers) to Succeed*. Tokyo: JALTCALL Publications.
- Burston, J., and Giannakou, K. (2021). MALL language learning outcomes: a comprehensive meta-analysis 1994–2019. *ReCALL J.* 34, 1–22. doi: 10.1017/S0958344021000240
- Chambers, A. (2019). Towards the corpus revolution? bridging the research–practice gap. *Lang. Teach.* 52, 460–475. doi: 10.1017/S0261444819000089
- Clenton, J., and Booth, P. (eds) (2020). *Vocabulary and the Four Skills: Pedagogy, Practice, and Implications for Teaching Vocabulary*. Milton Park: Routledge, doi: 10.4324/9780429285400
- Cobb, T., and Laufer, B. (2021). The nuclear word family list: a list of the most frequent family members, including base and affixed words. *Lang. Learn.* 71, 834–871. doi: 10.1111/lang.12452
- Cohen, J. (1988). *Statistical Power Analysis for the Behavioral Sciences*. Milton Park: Routledge.
- Council of Europe (2001). *Common European Framework of Reference for Languages: Learning, Teaching, Assessment*. Cambridge: Cambridge University Press.
- Coxhead, A. (2000). A new academic word list. *TESOL Quarterly* 34, 213–238. doi: 10.2307/3587951
- Dodigovic, M., and Agustín-Llach, M. P. (2020). “Introduction to vocabulary-based needs analysis,” in *Vocabulary in Curriculum Planning: Needs, Strategies and Tools*, eds M. Dodigovic and M. P. Agustín-Llach (London: Palgrave Macmillan), doi: 10.1007/978-3-030-48663-1\_1
- EFL Technologies (2017). *NGSL Builder Multilingual (1.1)*. Available online at: <https://efltechnologies.com/our-products/android-apps/>. (accessed June 13, 2022).
- Gardner, D., and Davies, M. (2014). A new academic vocabulary list. *Appl. Ling.* 35, 305–327. doi: 10.1093/applin/amt015
- Hao, T., Wang, Z., and Ardasheva, Y. (2021). Technology-Assisted vocabulary learning for EFL learners: a meta-analysis. *J. Res. Educ. Effect.* 14, 645–667. doi: 10.1080/19345747.2021.1917028
- Lai, C., and Gu, M. (2011). Self-regulated out-of-class language learning with technology. *Comp. Assis. Lang. Learn.* 24, 317–335. doi: 10.1080/09588221.2011.568417
- Laufer, B. (2013). Lexical thresholds for reading comprehension: what they are and how they can be used for teaching purposes. *TESOL Quarterly* 47, 867–872. doi: 10.1002/tesq.140
- Li, Y., and Hafner, C. A. (2022). Mobile-assisted vocabulary learning: investigating receptive and productive vocabulary knowledge of Chinese EFL learners. *ReCALL* 34, 66–80. doi: 10.1017/S0958344021000161
- Lin, J.-J., and Lin, H. (2019). Mobile-assisted ESL/EFL vocabulary learning: a systematic review and meta-analysis. *Comp. Assisted Lang. Learn.* 32, 878–919. doi: 10.1080/09588221.2018.1541359
- Mahdi, H. S. (2017). Effectiveness of mobile devices on vocabulary learning: a meta-analysis. *J. Educ. Comp. Res.* 56, 134–154. doi: 10.1177/0735633117698826
- Nation, I. S. P. (2013). *Learning Vocabulary in Another Language*, 2nd Edn. Cambridge: Cambridge University Press, doi: 10.1017/CBO9781139858656
- Pallant, J. (2016). *SPSS Survival Manual: A step by step guide to data analysis using IBM SPSS*, 6th Edn. Milton Park: Routledge.
- Sage, K., Krebs, B., and Grove, R. (2019). Flip, slide, or swipe? learning outcomes from paper, computer, and tablet flashcards. *Technol. Knowledge Learn.* 24, 461–482. doi: 10.1007/s10758-017-9345-9
- Stockwell, G. (2013). “Technology and motivation in english-language teaching and learning,” in *International Perspectives on Motivation*, ed. E. Ushioda (London: Palgrave Macmillan), 156–175. doi: 10.1057/9781137000873\_9
- Stoeckel, T., and Bennett, P. (n.d.). *The New General Service List Test (NGSLST)*. Available online at: <http://www.newgeneralservicelist.org/ngsl-levels-test> (accessed March 22, 2020).
- Test Your English (2022). *Cambridge University Press & Assessment*. Available online at: <https://www.cambridgeenglish.org/test-your-english/general-english/> (accessed June 13, 2022).

- Webb, S., and Chang, A. C.-S. (2012). Second language vocabulary growth. *RELC J.* 43, 113–126. doi: 10.1177/0033688212439367
- Webb, S., and Nation, I. S. P. (2017). *How Vocabulary is Learned*. Oxford: Oxford University Press.
- Xodabande, I., and Atai, M. R. (2020). Using mobile applications for self-directed learning of academic vocabulary among university students. *Open Learn. J. Open Distance e-Learn.* 1–18. doi: 10.1080/02680513.2020.1847061 [Epub ahead of print].
- Xodabande, I., Pourhassan (Aydin), A., and Valizadeh, M. (2022). Self-directed learning of core vocabulary in English by EFL learners: comparing the outcomes from paper and mobile application flashcards. *J. Comp. Educ.* 9, 93–111. doi: 10.1007/S40692-021-00197-6
- Yang, X., Kuo, L.-J., Eslami, Z. R., and Moody, S. M. (2021). Theoretical trends of research on technology and L2 vocabulary learning: a systematic review. *J. Comp. Educ.* 8, 465–483. doi: 10.1007/s40692-021-00187-8
- Yu, A., and Trainin, G. (2021). A meta-analysis examining technology-assisted L2 vocabulary learning. *ReCALL FirstView* 34, 1–18. doi: 10.1017/S0958344021000239
- Yüksel, H. G., Mercanoğlu, H. G., and Yılmaz, M. B. (2020). Digital flashcards vs. wordlists for learning technical vocabulary. *Comp. Assisted Lang. Learn.* 1–17 [Epub ahead of print].
- Conflict of Interest:** The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.
- Publisher's Note:** All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.
- Copyright © 2022 Rahmani, Asadi and Xodabande. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.



# A Corpus Study of Lexical Bundles Used Differently in Dissertations Abstracts Produced by Chinese and American PhD Students of Linguistics

Kai Bao<sup>1</sup> and Meihua Liu<sup>2\*</sup>

<sup>1</sup> College of Foreign Languages, University of Shanghai for Science and Technology, Shanghai, China, <sup>2</sup> Department of Foreign Languages and Literatures, Tsinghua University, Beijing, China

## OPEN ACCESS

### Edited by:

Xuesong Gao,  
University of New South Wales,  
Australia

### Reviewed by:

Liang Li,  
University of Waikato, New Zealand  
Ismael Xodabande,  
Kharazmi University, Iran

### \*Correspondence:

Meihua Liu  
liumeihua@mail.tsinghua.edu.cn;  
ellenlmh@yahoo.com

### Specialty section:

This article was submitted to  
Language Sciences,  
a section of the journal  
Frontiers in Psychology

**Received:** 10 March 2022

**Accepted:** 06 June 2022

**Published:** 30 June 2022

### Citation:

Bao K and Liu M (2022) A Corpus  
Study of Lexical Bundles Used  
Differently in Dissertations Abstracts  
Produced by Chinese and American  
PhD Students of Linguistics.  
Front. Psychol. 13:893773.  
doi: 10.3389/fpsyg.2022.893773

This study examined lexical bundles (LBs) used differently by Chinese and American PhD students of linguistics in their dissertation abstracts. Two corpora were built, with each having 700 dissertation abstracts produced by Chinese and American PhD students of linguistics, respectively. The study then used lexical analysis software to retrieve frequently used three-word LBs, from which LBs having different frequencies at a significant level across the two corpora were identified and termed as bundles used differently (BUDs). BUDs were then categorized and analyzed manually in terms of structure, function, and distribution in rhetorical moves. The major findings were: (1) 57.14% of the frequently used LBs were BUDs, of which 90.67% had occurrences in both corpora. The BUDs distributed inequivalently across categories and moves, with the text-oriented category and the move of Result having the most BUDs; and (2) BUDs exhibited two major patterns: the Chinese and American students filled different constituents into structurally and functionally similar constructions, and used LBs of dissimilar functions to fulfill the same communicative purposes. These findings indicate that variations in LB use have a high pedagogic value and confirm the need for using corpora to identify and teach core genre-specific vocabularies to second/foreign language learners.

**Keywords:** lexical bundle, dissertation abstract, corpora, rhetorical move, structure of BUDs, function of BUDs, distribution of BUDs

## INTRODUCTION

Formulaic language has received increasing attention from linguistics researchers over the past 3 decades. Despite an inconsistency of terminologies, such as constructions (Fillmore, 1988), idiomatic expressions (Titone and Connine, 1999), and LBs (Biber et al., 1999), a consensus has been reached that a substantial proportion of our language in use is comprised by formulaic language (Schmitt and Carter, 2004). Research shows that “stored and retrieved whole” (Wray, 2002, p. 9) multi-word items, an important type of formulaic language, account for as much as 58.60% of English-spoken discourse and 52.30% of written discourse (Erman and Warren, 2000), and are, therefore, “important building blocks of discourse in spoken and written registers”

(Biber and Barbieri, 2007, p. 263). Research reveals that learners process formulaic sequences in a considerably shorter reaction time than they process non-formulaic ones in reading (Jiang and Nekrasova, 2007) and speaking (Lin, 2010).

Formulaic language competence is increasingly considered as an indicator of language proficiency (Cortes, 2004). In academic prose, the mastery of discipline-specific sequences demonstrates membership of a discourse community (Segalowitz, 2010; Ädel and Erman, 2012). Formulaic language is particularly important for L2 (second language) learners because “it reduces the learning burden while maximizing communicative ability” (Ellis, 1994, p. 86). Learning formulaic sequences can effectively promote learners’ genre competence by offering ready-made sets of items to work with (Myles, 2004; Coxhead and Byrd, 2007), and allows them to use language with “a single mental effort” (Hunston, 2002, p. 174). Formulaic language, therefore, has high pedagogic value for English for academic purpose (EAP) instruction for L2 learners.

Lexical bundles (LBs), as an important component of formulaic language, are “sequences of words that commonly go together in natural discourse” (Biber and Conrad, 1999, p. 184). The defining features of LBs are non-idiomaticity, structural incompleteness, and frequency-driven identification, with which LB research extends to linguistic items rarely captured with conventional approaches. For instance, *in this dissertation I* is a four-word LB used by American PhD students to present goals of research projects. The sequence, however, is rarely identified by traditional phraseology due to its grammatical incompleteness.

LBs have been widely examined across EAP contexts, including textbooks (e.g., Biber et al., 2004; Liu and Chen, 2020), research papers across different disciplines (e.g., Cortes, 2004; Hyland, 2008a,b), research papers or essays produced by L1 (first language) and L2 English writers (e.g., Chen and Baker, 2010; Pan et al., 2016; Bychkovska and Lee, 2017), and so on. LBs in abstracts have been examined mainly in L1 and L2 English writers’ research articles (RAs) (e.g., Hu and Huang, 2017; Liu and Lu, 2019; Kim and Lee, 2021), whereas those in other genres, such as conference (e.g., Wongwiwat, 2016) and dissertation abstracts (e.g., Lu and Deng, 2019), have been inadequately researched. Dissertations and RAs are distinct genres and use language differently (Hyland, 2008a; El-Dakhs, 2020). Coupled with the rapidly rising number of graduate students who are going to be professionals and need to write in English in different areas in China (e.g., the number of new graduate students was over 1.1 million in 2021, while it was around 0.5 million in 2016), use of LBs in dissertation abstracts, which are the only section required to be written in English as well in most disciplines, deserves research.

In addition, the current LB literature approaches different use of LBs primarily by comparing LBs beyond a frequency cut-off, but rarely by directly examining those having significantly different frequencies across corpora. For instance, Dahunsi and Ewata (2022) compared three-word LBs having 50 or more occurrences in L1 and L2 writers’ texts, showing different structural patterns and colligational characteristics. Their results revealed certain features of LBs used by L1 and L2 writers (e.g., L2 writers used a greater number/proportion of LBs within a certain

structure than L1 writers), but did not comprehensively reveal what LBs were used with frequencies different at a significant level by the two groups. It is often assumed that shared bundles in L1 and L2 corpora have often been learned or acquired by L2 learners, while the bundles used (significantly) differently often indicate L2 learners’ tendency to use the bundles that they have accustomed/trained to and inability to freely employ the bundles frequently used by L1 speakers. Thus, LBs used significantly differently by L1 and L2 learners are more useful to L2 learners in that they represent items to be specifically noticed in L2 learners’ academic writing to enhance its genre-nativeness. Consequently, the present research seeks to examine the LBs used differently by Chinese and American PhD students of linguistics in their dissertation abstracts, hoping to help teach and learn this specific genre more effectively.

## STUDIES ON LEXICAL BUNDLES

LBs are “recurrent expressions, regardless of their idiomaticity, and regardless of their structural status” (Biber et al., 1999, p. 990). LBs have transparent meanings interpretable by their individual components (Liu and Chen, 2020), and, therefore, perform categorical discourse functions (Oakey, 2020) and fulfill particular pragmatic purposes (Biber et al., 2004).

LB research often uses pre-defined taxonomies to categorize and analyze LBs in terms of structure, function, as well as communicative purpose. Popular structural (Table 1) and functional taxonomies (Table 2) are proposed by Biber et al. (1999) and Hyland (2008a) for LBs in academic discourse, respectively. In addition, Swales and Peak’s (2009) five-move model (Table 3) has been used in research on LBs in abstracts. LB researchers generally adopt/adapt these models according to the scope and need of their studies.

Comparative LB research, when at its birth, often compared LBs beyond a frequency cut-off across corpora. For example, Biber et al. (1999) conducted the first LB research and found that LBs in conversation mainly consisted of clause segments,

**TABLE 1** | Biber et al.’s (1999, p. 1014–1015) structural taxonomy.

Structure	Example
Noun phrase + of	<i>the nature of the, a large number of</i>
Noun phrase + other post-modifier fragment	<i>the way in which, the relationship between</i>
Prepositional phrase + of	<i>on the basis of, in the context of</i>
Other prepositional phrase	<i>on the other hand, with respect to the</i>
Be + noun/adjective phrase	<i>is due to the, is the same as</i>
Passive verb + prepositional phrase fragment	<i>is based on the, can be found in</i>
Anticipatory <i>it</i> + verb/adjective phrase	<i>it is important to, it should be noted</i>
(Verb phrase) + <i>that</i> -clause fragment	<i>should be noted that, we assume that the</i>
(Verb/adjective) + <i>to</i> -clause fragment	<i>to be able to, are likely to be</i>
Adverbial clause fragment	<i>if there is a, as can be seen in</i>
Pronoun/noun phrase + <i>be</i>	<i>this is not the, this is the first</i>
Other expressions	<i>than that of the, may or may not</i>



**TABLE 2 |** (Hyland, 2008a, p. 49) functional taxonomy.

Category	Examples
<b>Research-oriented:</b> help writers to structure their activities and experiences of the real world	
Location: indicate time and place	<i>at the beginning of, at the same time</i>
Procedure	<i>the use of the, the operation of the</i>
Quantification	<i>the magnitude of the, a wide range of</i>
Description	<i>the structure of the, the size of the</i>
Topic: related to the field of research	<i>in the Hong Kong, the currency board system</i>
<b>Text-oriented:</b> concern with the organization of the text and the meaning of its elements as a message or argument.	
Transition signals: establishing additive or contrastive links between elements	<i>on the other hand, in addition to the, in contrast to the</i>
Resultative signals: mark inferential or causative relations between elements	<i>as a result of, it was found that, these results suggest that</i>
Structuring signals: text-reflexive markers which organize stretches of discourse or direct reader elsewhere in text	<i>in the present study, in the next section</i>
Framing signals: situate arguments by specifying limiting conditions	<i>in the case of, with respect to the, on the basis of</i>
<b>Participant-oriented:</b> focus on the writer or reader of the text.	
Stance features: convey the writer's attitudes and evaluations	<i>are likely to be, may be due to, it is possible that</i>
Engagement features: address readers directly	<i>it should be noted that, as can be seen</i>

*The bold font denotes major categories.*

including declarative structures (e.g., *I don't know what*) and interrogative structures (e.g., *what are your doing*), and that those in academic prose mainly consisted of nominal (e.g., *the end of the*) or prepositional elements (e.g., *on the basis of*). Later, LB studies began to compare LBs frequently used across corpora, featuring different registers (e.g., Biber, 2006; Biber and Barbieri, 2007; Huang, 2018), disciplines (e.g., Cortes, 2004; Hyland, 2008b; Liu and Chen, 2020), genres (e.g., Cortes, 2004; Hyland, 2008a; Gao, 2017), and writer groups (e.g., Nekrasova, 2009; Chen and Baker, 2010; Ädel and Erman, 2012).

Most research on abstracts focuses on LBs in RAs across rhetorical moves (e.g., Omidian et al., 2018; Abdollahpour and Gholami, 2019; Qi and Pan, 2020) and written by L1 and L2 English writers (e.g., Niu, 2014; Hu, 2015; Hu and Huang, 2017), revealing distributional patterns of LBs used differently across corpora. Several recent studies have examined LBs in dissertation abstracts. Lu and Deng (2019) compared LBs in 13,596 and 4,755 dissertation abstracts produced by PhD students of all available disciplines at Tsinghua University and Massachusetts Institute of Technology, respectively. The results showed substantial frequency differences of LBs in structural and functional categories (e.g., the Chinese and American students used 1,640.83 and 786.87 tokens of VP-based LBs per million words, respectively). Similar results were achieved by

Lyu and Gee (2020) who compared thesis abstracts written by Chinese and American master students of culture, linguistics, literature, and pedagogy. Li et al. (2020) conducted a bundle-driven analysis to reveal sentence initial LBs in rhetorical moves of dissertation abstracts produced by British PhD students of arts and humanities, social sciences, life sciences, and physics. The study identified a new rhetorical move—structure. Even so, research on LBs in dissertation abstracts is still limited.

Probably because of complexity, little research has specially investigated LBs used with frequencies different at a significant level across corpora, which helps reveal what and how certain LBs are used differently across corpora. Hence, what LBs are used differently for the same functions and communicative purposes is still inadequately researched. This study defines LBs used differently as items whose frequencies are different at a significant level ( $p < 0.05$ ) across corpora, as shown by the log-likelihood (LL) value equal to or greater than 3.84. LL tests perform effectively particularly in comparing the frequencies of low-frequency items across corpora of different sizes (Dunning, 1993; Rayson and Garside, 2000), and are considered “useful for comparing the relative frequency of words or phrases” (Simpson-Vlach and Ellis, 2010, p. 492). Despite its limitation suggested by Bestgen (2017), the LL statistics has been widely used in LB or n-gram studies (e.g., Simpson-Vlach and Ellis, 2010; Ädel and Erman, 2012; Du, 2013; Pan et al., 2016; Bychkovska and Lee, 2017; Hu and Huang, 2017; Hyland and Jiang, 2018, 2021; Lu and Deng, 2019; Jiang and Hyland, 2022).

The LBs used differently by L1 and L2 English writers often indicate L2 learners' tendency to use the bundles that they have accustomed/trained to and inability to freely employ the bundles frequently used by L1 speakers. Thus, knowledge of differently used bundles is useful to instructors who can use the knowledge to formulate an appropriate pedagogy to help L2 learners use a more variety of LBs in their academic writing and thus enhance

**TABLE 3 |** (Swales and Feak, 2009, p. 5) five-move model of abstracts.

Move #	Primary label	Primary function
Move 1	Background	Introducing background/situation
Move 2	Goal	Presenting research/aim
Move 3	Methodology	Stating methodology/materials/subjects/procedures
Move 4	Result	Displaying results/findings
Move 5	Conclusion	Discussing conclusions/implications

its genre-nativeness. English has become “global lingua franca of academia” (Mauranen et al., 2010, p. 183) that “is no one’s first language” (Hyland, 2019, p. 19). Hence, appropriate use of commonly used formulaic English in dissertation abstracts can help promote Chinese PhD students’ readership and academic influence in the international academic community (Hu and Huang, 2021).

All these motivated the present research, which aimed to identify the profiles of LBs used differently at a significant level by Chinese and American PhD students in their linguistics dissertation abstracts and examine how they are used differently by the two groups. And the following two research questions were formulated:

- (1) What are the profiles of LBs used differently in dissertation abstracts produced by Chinese and American PhD students of linguistics?
  - a. What are the forms and frequencies of the LBs used differently in the two corpora?
  - b. What are the functional and rhetorical move distributions of the LBs used differently in the two corpora?
- (2) How do the two groups use different LBs to achieve the same functions and communicative purposes in their dissertation abstracts?
  - a. How are LBs used for the same functions in the two corpora?
  - b. How are LBs used for the same communicative purposes in the two corpora?

## RESEARCH DESIGN

### Corpora

We first searched for dissertations completed by Chinese PhD students of linguistics between 2000 and 2020 from 13 Chinese universities via China National Knowledge Infrastructure, the National Library of China, and relevant university libraries. All the universities were qualified to reward doctoral degrees and tier-1 universities recognized by the Ministry of Education of the People’s Republic of China. The time frame covered a substantial proportion of the genre given that many of the PhD programs in China were founded around 2000. Finally, we gathered 700 dissertation abstracts and established the Chinese University Linguistics Dissertation Abstracts Collection (CUC), which had an overall token of 613,713 and a mean token of 876.73 (SD = 509.44).

To parallel with Chinese universities, high-ranking American universities on 2020 QS World University Rankings by linguistics were selected. Seven hundred dissertations abstracts completed by PhD students of linguistics from these 13 universities in the same time frame were gathered via ProQuest and university libraries to establish the American University Linguistics Dissertation Abstracts Collection (AUC). It had an overall token of 247,359 and a mean token of 353.37 (SD = 124.63). The source universities and amounts of the sample texts are presented in the **Supplementary Material**.

## Data Analyzing Framework

### Identifying LBs

The present research focused on three-word LBs because they are considerably more frequent than four- and five-word LBs (Hyland, 2008a) and the focus of many existing LB studies (e.g., Hu, 2015; Huang, 2018; Azad and Khiabani, 2018). A consistent selection of length promotes comparability. This study used the WordSmith Tools 8.0 (Scott, 2020) to retrieve three-word LBs based on a cut-off point of 60 occurrences per million words (pmw) in at least 2% (14 texts in CUC and AUC, respectively) of the sample texts. This study adopted a relatively higher frequency cut-off point because the dispersion rate is a more salient parameter in research concerning abstracts that are generally mini texts. In this study, for instance, an adoption of Hyland (2008a,b) 10% dispersion rate entailed 283 occurrences pmw for AUC, and 114 for CUC, which yielded less than 30 items. We, therefore, adopted 2% dispersion cut-off, which was 14 texts and thus 60 occurrences pmw. The cut-off point was consistent with and even stricter than Lu and Deng’s (2019) 20 occurrences and 2.70–9.70‰, and Omidian et al.’s (2018) 20 occurrences and 5.10‰ parameters.

After the retrieval, the overlapping items were combined with the approach proposed by Chen and Baker (2010) to reduce frequency inflation. The LBs underlying complete subsumption where two or more LBs overlapped and one LB subsumed the other(s) were combined (e.g., *as well as* subsumed all occurrences of *well as the* and *well as a* that were, therefore, combined into *as well as*). Those underlying complete overlaps where two three-word LBs were actually one four-word LBs were combined into longer items (e.g., all tokens of *it is argued* and *is argued that* were largely similar and, therefore, were combined into *it is argued that*).

To identify LBs whose frequencies varied at a significant level across CUC and AUC, we used Rayson’s (2016) Effect Size Calculator to yield the LL values. LBs of LL values greater than 3.84 ( $p < 0.05$ ) were recognized as bundles used differently (BUDs). BUDs of a significantly greater frequency in CUC were identified as CUC BUDs and those in AUC as AUC BUDs. We also used the Effect Size Calculator to identify LBs of a significantly greater frequency in a rhetorical move than in a whole text. LBs with LL values greater than 3.84 ( $p < 0.05$ ) were recognized as move-specific bundles and considered a key to constructing a specific move, given their specific occurrences.

### Categorizing LBs

This study categorized LBs in terms of their structure, function, and distribution in rhetorical moves. The working structural taxonomy, functional taxonomy, and the rhetorical moves, as well as communicative purposes, are provided in the **Supplementary Material**.

Structurally, we used Biber et al.’s (1999) taxonomy developed specifically for academic prose, and conducted modifications to enhance its relevance to the current genre. For instance, we replaced the *noun + verb phrases + that-clause* with *subject + verb phrases + (that-clause)* to account for a considerable number of AUC items containing first person singular *I*.

Functionally, we used Hyland's (2008a) functional taxonomy developed specifically for LBs in RAs and dissertations, and added new sub-categories, including relationship signals (LBs denoting relationships between entities, such as *is consistent with*), objective signals (LBs denoting objectives, such as *in order to*), and other bundles. We also promoted inferential and causative signals to sub-categories within the text-oriented category. To reduce subjectivity on functional categorization, a PhD student of linguistics was employed to categorize LBs independently with the researcher. Our Cohen's Kappa coefficient was 0.889 ( $k > 0.80$ ), indicating an excellent strength of inter-coder agreement (Landis and Koch, 1977). For LBs whose functions were disagreed upon, the two coders went through and discussed the items until they reached complete consensus.

With regard to rhetorical moves, we used Swales and Feak's (2009) five-move model and followed Li et al. (2020) to add the move of structure that specifically outlined dissertation structures found in many CUC and AUC texts in the present study. Based on the model, the two researchers identified moves and their communicative purposes independently, with a Kappa coefficient of 0.813. We then went through and discussed the disagreed-upon items until we reached complete consensus.

### Analyzing LBs

As discussed, we identified LBs frequently used in the two corpora via WordSmith Tools 8.0 (Scott, 2020), with a cut-off point of 60 occurrences pmw in at least 2% of the sample texts. We then identified BUDs from the LBs frequently used by their LL value (LL value  $\geq 3.84$ ,  $p < 0.05$ ), with Rayson's (2016) Effect Size Calculator.

To reveal how LBs were used differently in the two corpora, we examined the structures of BUDs in the same functional categories to determine whether the BUDs of the same functions resulted from different structures. Instead of analyzing tokens of BUDs serving different functions in rhetorical moves, we analyzed the functions of BUDs identified as move-specific bundles, serving the same communicative purposes in rhetorical moves, because: (1) rhetorical moves in dissertation abstracts are often more informative than those in RAs in that dissertation abstracts are considerably longer and contain LBs, serving various communicative purposes; and (2) move-specific bundles are key items used to construct moves whose occurrences have move-specific features.

## RESULTS

### Profiles of Bundles Used Differently in the Two Corpora Forms and Frequencies

According to the cut-off point, 274 bundle types with 20,154 tokens and 195 types with 5,798 tokens were identified in CUC and AUC, respectively, as shown in Table 4 (full lists are in the Supplementary Material). In terms of token frequency, the Chinese students used a substantially greater number of LBs than their American counterparts (LL value = 543.41,  $p < 0.0001$ ). Regarding bundle forms, only 101 bundle types were shared in the

two corpora, suggesting that 63.14% of the bundle types frequent in CUC were not frequent in AUC, and that 48.21% of the bundle types frequent in AUC were not frequent in CUC. These results indicated substantial variations across CUC and AUC with regard to both frequencies and forms. The LBs frequently used by the Chinese and American students were, therefore, considerably dissimilar.

Of the frequently used LBs, 268 BUDs that represented the LBs used differently by the two groups were identified. Of the 268 BUDs, 175 were CUC BUDs, indicating that 63.87% of the LBs frequently used in CUC had significantly greater frequencies than those in AUC; 93 were AUC BUDs, indicating that 47.69% of those in AUC had significantly greater frequencies than those in CUC. It should be noted that 157 of the 175 (89.71%) CUC BUDs had occurrences in both collections, and the proportion was 86 of 93 (92.47%) for AUC BUDs. These results showed that the Chinese and American PhD students of linguistics might draw on a quite similar formulaic language resource but with substantially dissimilar frequencies.

### Distribution in Functional Categories and Rhetorical Moves

As shown in Table 5, the greatest number of BUDs occurred in the text-oriented category, followed by the research- and participant-oriented categories, and description bundles had the largest number of BUDs, followed by framing and structuring signals. It should be noted that inferential signals had 15 AUC, yet only three CUC BUDs, and that inferential and causative signals were the only two categories that had more AUC than CUC BUDs. Meanwhile, topic bundles had 14 CUC BUDs but no AUC BUDs; transition signals had seven CUC BUDs but only one AUC BUD. These results revealed an inequivalent distribution of BUDs across all the three functional categories and their 15 subcategories.

Table 6 shows that the greatest number of BUD tokens occurred within the move of Result, followed by moves of Methodology and Goal. These results suggested that the Chinese and American students used substantially different LBs to present results, designs, and purposes of their research projects. It should be noted that the moves of Result and Goal were the only two rhetorical moves that had a greater proportion of AUC than CUC BUDs, while the move of Background had 16.16% of CUC yet only 7.27% of AUC BUDs. These findings indicated the Chinese students' heavy reliance on formulaic language in composing their dissertation abstracts.

### Analyses of BUDs in the Two Corpora BUDs Within Functional Categories

Figure 1 presents the structural distribution of BUDs in the three major functional categories. As shown in Figure 1A, the CUC BUDs fell into ten subcategories, whereas the AUC BUDs fell into five categories. Four structural categories, including *noun phrase + of*, *noun phrase with other post modifier fragments*, *other noun phrase fragment*, and *pronoun/noun phrase + be*, contained both CUC and AUC BUDs and accounted for 84.13% of all BUD tokens within the research-oriented category. The structural categories specific to AUC and CUC BUDs, including

**TABLE 4 |** Number of frequently used LBs and BUDs.

Corpus	LBs		BUDs		BUD occurrences in CUC and AUC	
	Type	Token	Type	Percentage in LBs	Type	Percentage in BUDs
CUC	274	20,154	175	63.87%	157	89.71%
AUC	195	5,798	93	47.69%	86	92.47%

**TABLE 5 |** Functional distribution of BUDs.

Category	CUC BUDs		AUC BUDs		Total	
	Type	Percentage	Type	Percentage	Type	Percentage
<b>Research-oriented</b>	<b>71</b>	<b>40.57%</b>	<b>31</b>	<b>33.33%</b>	102	<b>38.06%</b>
Location	3	1.71%	1	1.08%	4	1.49%
Procedure	14	8%	6	6.45%	20	7.46%
Quantification	8	4.57%	8	8.60%	16	5.97%
Description	32	18.29%	16	17.20%	48	17.91%
Topic	14	8%	0	0	14	5.22%
<b>Text-oriented</b>	<b>94</b>	<b>53.71%</b>	<b>59</b>	<b>63.44%</b>	153	<b>57.09%</b>
Transition signals	7	4%	1	1.08%	8	2.99%
Inferential signals	3	1.71%	15	16.13%	18	6.72%
Causative signals	8	4.57%	10	10.75%	18	6.72%
Structuring signals	23	13.14%	11	11.83%	34	12.69%
Framing signals	23	13.14%	12	12.90%	35	13.06%
Relationship signals	16	9.14%	5	5.38%	21	7.84%
Objective signals	14	8%	5	5.38%	19	7.09%
<b>Participant-oriented</b>	<b>10</b>	<b>5.71%</b>	<b>2</b>	<b>2.15%</b>	12	<b>4.48%</b>
Stance features	5	2.86%	1	1.08%	6	2.24%
Engagement features	1	0.57%	1	1.08%	2	0.75%
Other functions	4	2.29%	1	1.08%	5	1.87%
Total	175	100%	93	100%	268	100%

The bold font denotes major categories and is not counted in the total number.

*subject + verb phrase + (that-clause)* and *other prepositional phrase*, etc., only accounted for 15.87% of all BUD tokens within this category. **Figure 1B** show that the CUC BUDs fell into all the 17 subcategories and the AUC BUDs fell into 13 subcategories, which accounted for 84.02% of BUDs within the text-oriented category. **Figure 1C** shows that the participant-oriented category had three LB types, and that only one structural category, *noun phrase + of*, was shared by CUC and AUC BUDs within this category and accounted for 45.70% of BUDs.

**TABLE 6 |** Rhetorical move distribution of BUDs.

Move	CUC BUDs		AUC BUDs	
	Token	Percentage	Token	Percentage
Background	2,203	16.16%	216	7.27%
Goal	1,714	12.57%	623	20.97%
Methodology	2,735	20.06%	560	18.85%
Result	4,777	35.04%	1,330	44.77%
Conclusion	1,825	13.39%	166	5.59%
Structure	378	2.77%	76	2.56%
Total	13,632	100%	2,971	100%

These results suggested that BUDs in the two corpora fell into substantially similar structural categories, meaning that what led to the two groups' different use of LBs with respect to structure was mainly different constituents filled in the functionally and structurally similar constructions, but not different structures used to construct LBs of the same functions.

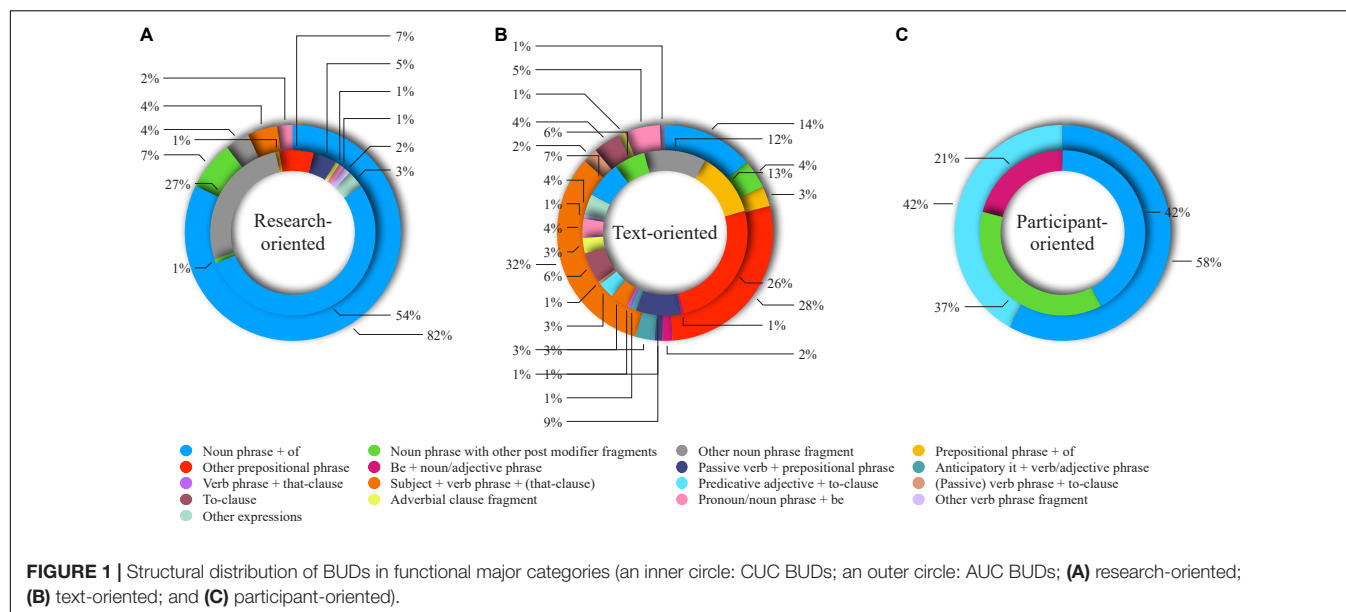
We then analyzed the constituents filled into the same structures of BUDs within the same functional categories. **Table 7** shows five typical pairs of CUC and AUC BUDs, which had the same structures and functions. For example, within location bundles, the Chinese students were more likely to use (*in*) *the process of*, whereas the American students used (*over*) *the course of* to refer to a time period. In Ex. 1 and Ex. 2, both BUDs collocated with *years* in similar contexts.

Ex. 1. In **the process of** 30 years from 1990 to the present, ... (CUC).

Ex. 2. The fieldwork for this study was conducted over **the course of** 7 years. (AUC).

Likewise, within description bundles, while the Chinese students tended to refer to the meaning of lexical items by *the meaning of*, their American counterparts more often used *the semantics of*. The two BUDs usually occurred in the same contexts, as shown in Ex. 3 and 4. The same pattern was





identified with many other pairs of BUDs within description bundles commonly used by the Chinese and American students, respectively, including *features of the* and *properties of the*, *characteristics of the* and *the nature of*, and *the framework of* and *a model of*.

Ex. 3. . . a “semantic mismatch” between **the meaning** of the lexical items. . . (CUC).

Ex. 4. . . a single underlying difference in **the semantics** of these lexical items. . . (AUC).

Analyses showed that, within inferential signals, the Chinese students used more *find* as in *it is found*, but the American students used more *argue* and *show* as in *it is argued/shown* in *anticipatory it + verb/adjective phrase* structure. These results indicated the Chinese students’ particular and even exclusive preference for using *find* to report research results. In contrast, the American students used a much wider range of verbs, including *argue*, *show*, *propose*, *suggest*, *indicate*, *demonstrate*, and *propose*, whose occurrences in CUC inferential signals were all significantly fewer. Despite the semantic differences of these verbs, *find* might be the first and most prototypical reporting verb learned and used by many Chinese students who thus felt more confident with the verb (Li et al., 2018). Within inferential signals of *subject + verb phrase + (that-clause)* structure, the Chinese students primarily used *we find that* and strictly avoided using LBs containing the first person singular *I* to reduce authorial stance, while these bundles were the most common inferential signals used by the American students. Table 8 shows a total of 10 AUC LBs containing authorial *I*, which were all AUC BUDs. These items had 1,827 tokens pmw in AUC, yet only 21 tokens in CUC. They accounted for 31.51% of the total AUC bundle tokens and were, therefore, the most important bundle group in AUC in this regard. The tokens of the most common *I*-bundle *I argue that* alone made up 3.24% of all bundle tokens in AUC. Functionally, the bundles fell into inferential signals and procedure bundles that indicated writers’ intellectual aspects in research. Meanwhile, 9 of the 10 items were move-specific bundles and thus a key to

the construction of specific rhetorical moves, including moves of AUC Result and Methodology. The *I*-bundles were, therefore, crucially important to the American students but strictly avoided by the Chinese students in the *subject + verb phrase + (that-clause)* structure.

Meanwhile, the Chinese students used significantly more structuring signals (LL value = 46.19,  $p < 0.0001$ ), such as the CUC BUD *the present study*, as sentence subjects. The subjects of move-specific bundles were primarily *this dissertation* as in *this dissertation examines* in AUC Goal and *the dissertation* as in *the dissertation is* in Structure. This finding showed the American students’ more flexible use of subjects to implicitly mark move boundaries and their attempts to highlight individual contributions to the design and results of their projects.

Within structuring signals, the American students consistently used *dissertation* as in *in this dissertation*, whereas the Chinese students used more often *study* as in *in this study*, *research* as *the present research*, and *dissertation* as in *the present dissertation* to refer to doctoral research. Many BUDs in this category exhibited this feature, including AUC BUDs *of the dissertation*, *this dissertation presents*, as well as *this dissertation is*, and

**TABLE 7 |** Example BUDs with different constituents in constructions.

Function	Structure	BUDs
Location bundle	Noun phrase + of	CUC: <i>the process of</i> AUC: <i>the course of</i>
Description bundle	Noun phrase + of	CUC: <i>the meaning of</i> AUC: <i>the semantics of</i>
Inferential signal	Subject + verb phrase + that-clause	CUC: <i>we find that</i> AUC: <i>I show that</i>
Causative signal	Noun phrase + of	CUC: <i>the findings of</i> AUC: <i>the results of</i>
Structuring signal	Other prepositional phrase	CUC: <i>in this study</i> AUC: <i>in this dissertation</i>

**TABLE 8 |** /-bundles.

/-bundle	Token frequency		Function	Move (communicative purpose)
	AUC	CUC		
<i>I argue that</i>	188	8	Inferential signal	Result (mark the report of results)
<i>I show that</i>	96	0	Inferential signal	Result (mark the report of results)
<i>I propose that</i>	40	2	Inferential signal	Result (mark the report of results)
<i>I demonstrate that</i>	22	0	Inferential signal	Result (mark the report of results)
<i>I focus on</i>	19	0	Procedure bundle	Methodology (display research acts)
<i>I examine the</i>	19	0	Procedure bundle	Methodology (display research acts)
<i>I propose a</i>	20	2	Inferential signal	NA
<i>I argue for</i>	18	1	Inferential signal	Result (mark the report of results)
<i>I suggest that</i>	15	0	Inferential signal	Result (mark the report of results)
<i>I show that</i>	15	0	Inferential signal	Result (mark the report of results)

the CUC BUDs *in this research*, *present study has*, and *this study is*. The difference might reflect the Chinese students' relative unfamiliarity with and, therefore, reluctance to use *dissertation* that specifically denoted doctoral research. The same pattern was identified within causative signals where the Chinese students tended to use *finding*, while the American students primarily relied on *result*, which was reflected in the CUC BUDs *the findings of* and *findings of the*, and the AUC BUD *the results of*. The Chinese students might use *findings* to emphasize their active efforts to *find* facts, whose subtle meaning could not be communicated by *result*. These findings indicate the two writer groups' different choices of words within the same structures.

### BUDs Within Rhetorical Moves

**Figure 2** presents the functional distribution of move-specific bundles in the communicative purposes that had most LBs in each of its six sub-moves. The communicative purposes included introducing previous studies (A), presenting purposes of research (B), displaying research acts (C), reporting results of research (D), demonstrating beneficiaries of research (E), and stating contents of chapters (F).

In Background, both groups used most LBs to achieve the communicative purpose of introducing previous studies where 11 BUDs were identified. As shown in **Figure 2A**, the Chinese students used a much wider functional variety of LBs that included description bundles (e.g., *the previous studies*), quantification bundles (e.g., *most of the*), framing signals (e.g., *the field of*), and procedure bundles (e.g., *the research on*). The American students only used framing signals (e.g., *the field of*) and causative signals (e.g., *it has been*). Similarly, to emphasize the importance of their research, the Chinese students used quantification bundles (e.g., *a lot of*), stance feature bundles (e.g., *the most important*), and transition signals (e.g., *is not only*), whereas the American students only used quantification bundles (e.g., *one of the*).

**Figure 2B** indicates both groups' marked reliance on objective signals to present research purposes in Goal. In CUC, 84.21% of the move-specific bundles were objective signals, which were 50% for AUC. This difference indicated the Chinese students' heavier reliance on objective signals to present research aims.

Instead, the American students showed marked reliance on structuring signals that made up 40% of the move-specific bundles, which were rarely used by the Chinese students. Ex. 5 and 6 show the two groups' typical use of objective and structuring signals in this regard, which shows that the American students' use of structuring signals commonly followed the pattern *this dissertation + investigates/examines/presents/explores*, while the Chinese students primarily relied on objective signals, such as *study aims to* that contained the verb *aim*, communicating a more explicit meaning of purposes.

Ex. 5. **This dissertation explores** the role of affect in sociolinguistic style. (AUC GOAL).

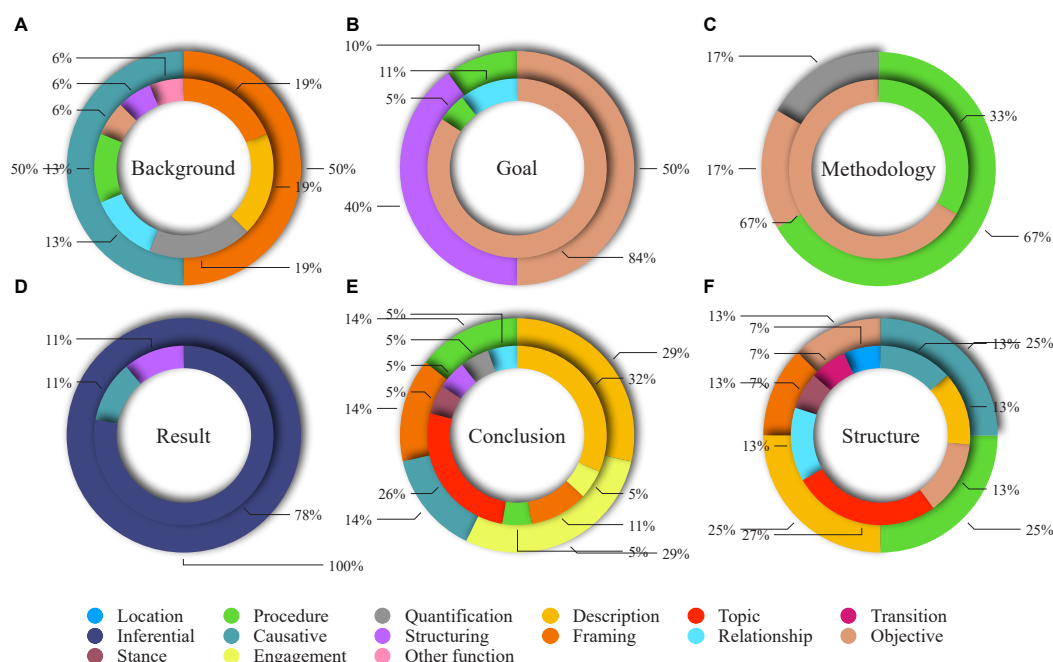
Ex. 6. The present **study aims to** explore the use of anaphora. (CUC GOAL).

The Chinese students' greater reliance on objective signals was also identified in displaying research acts in Methodology. **Figure 2C** shows that 66.67% of the CUC move-specific bundles fell into objective signals, which were only 16.67% for AUC. The American students used 66.67% of the items as procedure bundles that only accounted for 33.33% in CUC. The procedure bundles used by the American students commonly followed the pattern *I + examine/focus on the* in an active voice (Ex. 7). In comparison, objective signals, such as *to analyze the*, which fell into *to-clause* structure were commonly used by the Chinese students in a passive voice (Ex. 8). In addition, in description of research design in Methodology, the CUC items consisted of framing signals (55.56%), procedure bundles (22.22%), and quantification bundles (22.22%), while the AUC items fell equivalently into location bundles, framing signals, and quantification bundles. The Chinese students, therefore, used considerably more framing signals for the purpose. Ex. 9 is a series of CUC sentences that all began with a framing signal that specified the conditions underlying arguments. Such intense use of framing signals was not identified in any move of AUC.

Ex. 7. **I examine the** basic syntax of English comparatives and readers' expectations for.... (AUC).

Ex. 8. The Paired Sample *t*-test is conducted **to analyze the** tendency regarding. . .

Ex. 9. **In terms of** mode, we have. . . ; **in terms of** motive, we have. . . ; **in terms of** addressees, the offensive addressing. . . ; **in terms of** degree, we have. . . (CUC).



**FIGURE 2 |** Functional distribution of move-specific bundles in several communicative purposes (an inner circle: CUC; an outer circle: AUC; **(A)** introducing previous studies; **(B)** presenting purposes of research; **(C)** displaying research acts; **(D)** reporting results of research; **(E)** demonstrating beneficiaries of research; and **(F)** stating contents of chapters).

**Figure 2D** reveals both groups' marked reliance on inferential signals to signal the report of research results. It shows that 100% of the AUC move-specific bundles were inferential signals, which were 77.78% of the CUC items. In addition to objective signals, the CUC items included structuring and causative signals that could be a specific discourse signal used by the Chinese students for the purpose. The LL values of the CUC structuring signal *are as follows*, and causative signals *the major findings* were 34.05 and 17.73, respectively, but that of the inferential signal *we find that* was 13.32. The Chinese students specifically used the combination of *the major findings* and *are as follows* to signal the report of research results (Ex. 10). In comparison, the American students commonly used inferential signals in the pattern *I + argue/show/propose/demonstrate/suggest that*, which could highlight writers' individual contributions more explicitly (Ex. 11).

Ex. 10. **The major findings** of the present research **are as follows**: 1. . . 2. . . 3. . . (CUC RESULTS).

Ex. 11. **I show that** these two constructions again mirror the situation. . . **I propose that** the two constructions contribute the same semantic pieces. . . (AUC RESULTS).

Regarding Conclusion, **Figure 2E** reveals that, despite the two groups' similar use of description bundles (e.g., *the study of*) to point out the beneficiaries of their research, they generally used different categories as well as proportions of move-specific bundles for the purpose. It shows that 26.32% of the CUC items were topic bundles (e.g., *second language acquisition*) that specified the beneficiary field, which was not identified in AUC. The American students specifically used engagement feature bundles (e.g., *our understanding of*) to build relevance between readers and their research outcomes (Ex. 12). Using

participant-oriented bundles for the purpose as well, the Chinese students used the stance feature bundle *better understanding of* that also fell into *noun phrase + of* structure (Ex. 13). To mark the discussion of implications, the Chinese students used a combination of causative signals, structuring signals (Ex. 14), and procedure bundles, different from the American students' sole use of causative signals (Ex. 15).

Ex. 12. My work contributes to **our understanding of** how power. . . (AUC CONCLUSION).

Ex. 13. It is hoped that this study can contribute to a **better understanding of**. . . (CUC CONCLUSION).

Ex. 14. The present study also contributes to the pragmatic field **in the following** aspects. (1). . . (2). . . (3). . . (CUC CONCLUSION).

Ex. 15. **The implications of** this research extend to theories of discourse and reference. (AUC CONCLUSION).

**Figure 2F** shows the Chinese students' specific use of topic bundles, relationship signals, stance feature bundles, transition signals, as well as location bundles, and the American students' specific use of procedure bundles and framing signals to demonstrate the contents of dissertation chapters in Structure. Topic bundles accounted for 26.67% of the CUC items but none of the AUC items, which indicates the Chinese students' more recurrent mention of research subjects.

## DISCUSSION

### Profiles of BUDs in the Two Corpora

This study revealed that the Chinese students produced significantly more LBs than American students, consistent with

the finding in several previous studies that a considerably greater number of English LBs were used by L1 Chinese than L1 English writers (Guan and Zheng, 2005; Hyland, 2008a; Pang, 2009; Lou, 2010; Wei and Lei, 2011; Xu, 2012; Pan et al., 2016; Gao, 2017; Lyu and Gee, 2020). A similar finding is achieved with L1 Spanish speakers (Pérez-Llantada, 2014), L1 Iranian speakers (Jalali et al., 2008), and L1 Turkish speakers (Güngör and Uysal, 2016). These findings generally indicate L2 English writers' greater reliance on formulaic language to construct academic articles, as discussed in Hyland (2008a) and Paquot and Granger (2012). Hyland (2008a) argued that the considerably higher LB token frequency in L2 writers' academic writing primarily resulted from apprentice writers' heavier reliance on prefabricated items in the development of their arguments. Likewise, Paquot and Granger (2012, p. 139) argued that "less-proficient learners seem to be more reliant on lexical bundles." This is largely because L1 writers normally have a larger repertoire of formulaic language and thus do not have to stick to a smaller variety of high-frequency items, while L2 writers often have a smaller repertoire and have to use high-frequency LBs more recurrently.

In addition, the difference could be partially explained by the generic features of our corpora. Dissertation abstracts are a high-stake genre, concerning both research manifestation and degree fulfillment. While research manifestation is mainly for international readers, degree fulfillment needs to be approved by a dissertation supervision committee whose members are primarily Chinese professors. To cope with the complicated and important writing task in L2, Chinese students may tend to use language sequences with which they feel confident and specify situations of arguments more carefully, especially in a considerably longer text.

The present study also found that, although the LBs frequently used by the two groups were substantially different with respect to forms and frequencies, about 90% of the BUDs had occurrences in both collections. In addition, BUDs distributed inequivalently across functional categories and rhetorical moves. This finding was new since most existing LB studies focus mainly on frequently used LBs (e.g., Hyland, 2008a; Lu and Deng, 2019; Lyu and Gee, 2020), but rarely on LBs of significantly different frequencies across corpora. By focusing on BUDs, this study revealed that the Chinese and American students drew on a highly similar resource of formulaic sequences but selected items with considerably different frequencies. Nevertheless, these findings need to be confirmed in more similar research.

## Patterns of BUDs in the Two Corpora

This study showed that the Chinese and American students filled in different constituents into the structurally and functionally similar constructions, indicating the two groups' different choices of lexical bundles.

Within inferential signals of *subject + verb phrase + (that-clause)* structure, the Chinese students primarily used *we find that* and strictly avoided using LBs containing the first person singular *I* to reduce authorial stance, but these bundles were the most common inferential signals used by the American students. The first person plural refers to the research work or the author in CUC, which ought to be singular because

dissertations are individual works. The Chinese students' use of *we* probably resulted from their avoidance of authorial *I* or a "misunderstanding of rhetorical conventions" (Li et al., 2018, p. 42). Scholars in support of the avoidance of personal pronouns in academic prose argue that an unrestrained use of authorial *I* reduces the objective expression of ideas (Arnaudet and Barrett, 1984; Kirsch, 1994; Spencer and Arbon, 1996). The first person viewpoint, however, has become more acceptable in academia over recent years. Ivanic (1998) argued that the first person singular in academic writing is a powerful tool for self-representation and the construction of authorial identity. Similarly, Kuo (1999) and Hyland (2001) argued that a proper use of authorial *I* could promote the emphasis on the authors' individual contribution to a research field. Notably, Section 4.16 of the newest APA style (American Psychological Association, 2020) as well as Section 5.40 of the newest Chicago Manual of Style (The University of Chicago Press Editorial Staff, 2017) both encourage the use of first person pronouns to describe research work and personal reactions.

The study also revealed that the *I*-bundles were the most important AUC bundle group in terms of token frequency but were strictly avoided in CUC. The Chinese students' avoidance of using *I* might result from the conventional EAP protocol that requires writers to conceal the authorial aspects in favor of readers' closer focus on the substance of writing (Arnaudet and Barrett, 1984; Kirsch, 1994; Spencer and Arbon, 1996). Since dissertations might be the most important piece of writing in PhD students' lives (Hyland, 2008a), Chinese students are advised to use these bundles by substituting authorial *I* with names of research, chapters, or sections, such as *this dissertation/study/chapter/section argues that* for *I argue that*. The sequence is able to help Chinese students promote an authorial voice without breaking the protocol.

This study also showed that the Chinese and American students used LBs of dissimilar functions to fulfill the same communicative purposes. This might be related to the generic features of CUC. For instance, the causative signal *the major findings* and the structuring signal *are as follows* can be more effective discourse markers than inferential signals, such as *I show that* to signal a considerably longer report of results, as discussed in Allen (2009) and Qin (2014). The more frequent use of objective signals like *study aims to* in presenting research aims can be a more familiar discourse signal for Chinese professors, especially in a substantially longer text, so are structuring signals like *the present study* that marks implications in Conclusion, as discussed in Cortes (2004).

## CONCLUSION AND IMPLICATIONS

This study examined three-word LBs in 1,400 dissertation abstracts written by Chinese and American PhD students of linguistics and focused on BUDs, which were LBs used with significantly different frequencies by the two groups. The study had two major findings. First, most BUDs had occurrences in both corpora. The BUDs distributed inequivalently across functional categories and rhetorical moves, with the text-oriented



category and move of Result having the most BUDs. Second, the two groups used similar structures to construct LBs of the same functions but filled different constituents into structurally and functionally similar constructions. They also used dissimilar functions for the same communicative purposes in rhetorical moves. Many BUDs related to the considerably lengthier texts in CUC than in AUC. It should be noted that the Chinese students strictly avoided filling first person singular *I* in *subject + verb phrase + (that-clause)* structure to reduce authorial stance, while these bundles were the most common inferential signals used by the American students to report results of their research.

These findings reveal not only general profiles of BUDs in dissertation abstracts produced by Chinese and American PhD students of linguistics but also how they used BUDs to fulfill the same functions and communicative purposes. The general profiles indicate a similar lexical resource from which the Chinese and American students picked items with substantial different frequencies, categories, and moves of greater number of BUDs and thus greater variations. Hence, the bundle lists generated by this study can be used in teaching. An instructor can also use lexical analysis tools to identify bundles used differently by Chinese students compared with their American counterparts, and then formulate a bundle list. The instructor can then focus on reducing certain items and choosing a variety of appropriate items within a certain functional category or rhetorical move from her or his students' writing in order to enhance genre-nativeness. For instance, an instructor can ask his/her students to use more *(over) the course of* instead of *(in) the process of* to denote a time period, *the extent to which* to replace *the degree of*, and *specifically* to replace *that is the* in the same context. This can help reduce both CUC and AUC BUDs whose frequencies offset each other. Meanwhile, because of the availability of various corpora, students are encouraged to discover patterns of and differences in the use of LBs by different learner groups to improve their own writing.

Additionally, our finding of specific patterns underlying different uses of LBs shows that not all BUDs ought to be specifically noticed. The BUDs that relate to generic features (e.g., *of in the following* vs. *I show that*, the former marked a considerably lengthier move of Result in CUC), such as substantially different word counts and potentially different discursual expectations toward the genre, should not be the focus of instruction. Instead, the BUDs that relate to different linguistic choices in the same structure, function, as well as context should be paid closer attention to (e.g., both the *process of* vs. the *course of* denote a period of time). Instructors are also advised to teach the AUC BUDs containing authorial *I* to Chinese students by asking them to replace *I* with the name of study, chapter, or section (e.g., *I argue that* taught as *this dissertation argues that*). The revised sequence emphasizes individual contributions and, at the same time, adapts to the traditional EAP convention that discourages using authorial *I*. The instructor can also measure learning progress by using lexical analysis tools to compare LBs in L2 learners' writing against those produced by native learners of English. As discussed in Kazemi et al. (2014), bundle-based instruction raises

learners' awareness of formulaic language and develops their English writing skills.

This study confirms the need for using corpus linguistics to identify and teach core genre-specific vocabularies to L2 learners. Even so, the research has certain limitations. The biggest limitation is that it does not analyze the multifunctionality of LBs because of the subjective nature of functional categorization and the complexity of contexts. A study in a specific context may grant us a closer understanding of LB functions. Secondly, the study was based on corpora of texts with different average lengths that might affect the use of LBs (Pan et al., 2020). Future studies can use corpora of texts with a similar average length to complement our investigation. Thirdly, the present research only investigated dissertation abstracts produced by Chinese and American PhD students of linguistics. A comparative study between different disciplines will enrich the literature on dissertation abstracts. Such studies can help compile core genre-specific bundle lists that include structural, functional, and rhetorical move distributions to support instruction and research on EAP vocabularies. In addition, future studies can apply psycholinguistic experiments to determine the effects of and causes for BUDs. L1 and L2 English writers' reactions to BUDs can help instructors develop tailored teaching resources, techniques, and strategies.

## DATA AVAILABILITY STATEMENT

The raw data supporting the conclusions of this article will be made available by the authors, without undue reservation.

## ETHICS STATEMENT

Ethical review and approval was not required for the study on human participants in accordance with the local legislation and institutional requirements. Written informed consent for participation was not required for this study in accordance with the national legislation and the institutional requirements.

## AUTHOR CONTRIBUTIONS

KB and ML designed the research, analyzed the data, and wrote the manuscript. Both authors contributed to the article and approved the submitted version.

## ACKNOWLEDGMENTS

We would like to thank Shisheng Liu, the editors, and the reviewers for the constructive comments and suggestions.

## SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fpsyg.2022.893773/full#supplementary-material>

# REFERENCES

- Abdollahpour, Z., and Gholami, J. (2019). Embodiment of rhetorical moves in lexical bundles in abstracts of the medical sciences. *Southern Afri. Ling. Appl. Lang. Stud.* 37, 339–360. doi: 10.2989/16073614.2019.1692681
- Ädel, A., and Erman, B. (2012). Recurrent word combinations in academic writing by native and non-native speakers of English: a lexical bundles approach. *Eng. Specific Purposes* 31, 81–92. doi: 10.1016/j.esp.2011.08.004
- Allen, D. (2009). Lexical bundles in learner writing: an analysis of formulaic language in the ALESS learner corpus. *Komaba J. Eng. Educ.* 1, 105–127.
- American Psychological Association (2020). *Publication Manual of the American Psychological Association*, 7th Edn. Washington, DC: American Psychological Association.
- Arnauudet, M. L., and Barrett, M. E. (1984). *Approaches to Academic Reading And Writing*. Englewood Cliffs: Prentice Hall.
- Azad, H. R., and Khiabani, S. M. (2018). Lexical bundles in English abstracts of research articles written by Iranian scholars: examples from humanities. *Iran. J. Appl. Lang. Stud.* 10, 149–174.
- Bestgen, Y. (2017). Getting rid of the Chi-square and log-likelihood tests for analysing vocabulary differences between corpora. *Quad. Filol. Estud. Linguist.* 22, 33–56. doi: 10.7203/qf.22.11299
- Biber, D. (2006). *University Language: A Corpus-Based Study Of Spoken And Written Registers*. Amsterdam: John Benjamins. doi: 10.1075/sc1.23
- Biber, D., and Barbieri, F. (2007). Lexical bundles in university spoken and written registers. *Eng. Specific Purposes* 26, 263–286. doi: 10.1016/j.esp.2006.08.003
- Biber, D., and Conrad, S. (1999). “Lexical bundles in conversation and academic prose,” in *Out of Corpora: Studies in Honor of Stig Johansson*, eds H. Hasselgard and S. Oksefjell (Amsterdam: Rodopi), 181–189.
- Biber, D., Conrad, S., and Cortes, V. (2004). If you look at.: lexical bundles in university teaching and textbooks. *Appl. Ling.* 25, 371–405. doi: 10.1093/applin/25.3.371
- Biber, D., Johansson, S., Leech, G., Conrad, S., and Finegan, E. (1999). *Longman Grammar of Spoken and Written English*. Harlow, UK: Pearson Education.
- Bychkovska, T., and Lee, J. J. (2017). At the same time: lexical bundles in L1 and L2 university student argumentative writing. *J. Eng. Acad. Purposes* 30, 38–52. doi: 10.1016/j.jeap.2017.10.008
- Chen, Y., and Baker, P. (2010). Lexical bundles in L1 and L2 academic writing. *Lang. Learn. Technol.* 14, 30–49.
- Cortes, V. (2004). Lexical bundles in published and student disciplinary writing: examples from history and biology. *Eng. Specific Purposes* 23, 397–423. doi: 10.1016/j.esp.2003.12.001
- Coxhead, A., and Byrd, P. (2007). Preparing writing teachers to teach the vocabulary and grammar of academic prose. *J. Sec. Lang. Writ.* 16, 129–147. doi: 10.1016/j.jslw.2007.07.002
- Dahunsi, T. N., and Ewata, T. O. (2022). An exploration of the structural and colligational characteristics of lexical bundles in L1-L2 corpora for English language teaching. *Lang. Teach. Res.* doi: 10.1177/13621688211066572 [Epub ahead of print].
- Du, J. (2013). *The Use Of Lexical Bundles By Chinese EFL English-Major Undergraduates At Different University Levels: A Corpus-Based Study of L2 Learners’ Examination Essays Ph. D. Thesis*. The Chinese University of Hong Kong].
- Dunning, T. (1993). Accurate methods for the statistics of surprise and coincidence. *Comput. Ling.* 19, 61–74.
- El-Dakhs, D. A. S. (2020). Why are abstracts in PhD theses and research articles different? A genre-specific perspective. *J. Eng. Acad. Purposes* 36, 48–60. doi: 10.1016/j.jeap.2018.09.005
- Ellis, R. (1994). *The Study of Second Language Acquisition*. Oxford: Oxford University Press.
- Erman, B., and Warren, B. (2000). The idiom principle and the open-choice principle. *Text* 20, 29–62. doi: 10.1515/text.1.2000.20.1.29
- Fillmore, C. (1988). The mechanisms of construction grammar. *Berkeley Ling. Soc.* 14, 35–55. doi: 10.3765/bls.v14i0.1794
- Gao, X. (2017). A comparable-corpus-based study on native English and Chinese academic writers’ use of English lexical bundles. *For. Lang. Their Teach.* 3, 42–52.
- Guan, B., and Zheng, S. (2005). A corpus-based contrastive study of recurrent word combinations in English essays of Chinese college students and native speakers. *CELEA J.* 28, 37–48.
- Güngör, F., and Uysal, H. H. (2016). A comparative analysis of lexical bundles used by native and non-native scholars. *Eng. Lang. Teach.* 9, 176–188. doi: 10.5539/elt.v9n6p176
- Hu, X. (2015). A comparable study on lexical characteristics in different moves of English abstracts by Chinese and English writers. *Mod. For. Lang.* 38, 813–822.
- Hu, X., and Huang, Y. (2017). A contrastive study on the use of common high-frequency lexical bundles in different moves of English abstracts in Chinese and English engineering journals. *For. Lang. Learn. Theory Pract.* 4, 31–36.
- Hu, X., and Huang, Y. (2021). A contrastive study on interpersonal functions of shell nouns in different moves of English abstracts of Chinese and English scientific papers. *For. Lang. Learn. Theory Pract.* 3, 56–63.
- Huang, K. (2018). Register features of lexical bundles used by Chinese EFL majors: a contrastive analysis of spoken and written English. *For. Lang. World* 5, 71–79.
- Hunston, S. (2002). “Pattern grammar, language teaching, and linguistic variation: applications of a corpus-driven grammar,” in *Using Corpora to Explore Linguistic Variation*, eds R. Reppen, S. M. Fitzmaurice, and D. Biber (Amsterdam: John Benjamins), 167–183. doi: 10.1075/sc1.9.11hun
- Hyland, K. (2001). Humble servants of the discipline? Self-mention in research articles. *Eng. Specific Purposes* 20, 207–226. doi: 10.1016/S0889-4906(00)00012-0
- Hyland, K. (2008a). Academic clusters: text patterning in published and postgraduate writing. *Int. J. Appl. Ling.* 18, 41–62. doi: 10.1111/j.1473-4192.2008.00178.x
- Hyland, K. (2008b). As can be seen: lexical bundles and disciplinary variation. *Eng. Specific Purposes* 27, 4–21. doi: 10.1016/j.esp.2007.06.001
- Hyland, K. (2019). “Participation in publishing: the demoralizing discourse of disadvantage,” in *Novice Writers and Scholarly Publication: Authors, Mentors, Gatekeepers*, eds P. Habibie and K. Hyland (Palgrave Macmillan), 13–33. doi: 10.1007/978-3-319-95333-5\_2
- Hyland, K., and Jiang, F. (2021). Academic naming: changing patterns of noun use in research writing. *J. Eng. Ling.* 49, 255–282. doi: 10.1177/00754242211019080
- Hyland, K., and Jiang, F. K. (2018). Academic lexical bundles: how are they changing? *Int. J. Corpus Ling.* 23, 383–407. doi: 10.1075/ijcl.17080.hyl
- Ivanic, R. (1998). *Writing and Identity: The Discoursal Construction of Identity in Academic Writing*. Amsterdam: Benjamins. doi: 10.1075/swll.5
- Jalali, H., EslamiRasekh, A., and TavangarRizi, M. (2008). Lexical bundles and intradisciplinary variations: the case of applied linguistics. *Iran. J. Lang. Stud.* 2, 447–484.
- Jiang, F., and Hyland, K. (2022). COVID—19 in the news: the first 12 months. *Int. J. Appl. Ling.* 2022:12412. doi: 10.1111/ijal.12412
- Jiang, N., and Nekrasova, T. M. (2007). The processing of formulaic sequences by second language speakers. *Mod. Lang. J.* 91, 433–445. doi: 10.1111/j.1540-4781.2007.00589.x
- Kazemi, M., Katiraei, S., and Rasekh, A. E. (2014). The impact of teaching lexical bundles on improving Iranian EFL students’ writing skill. *Procedia Soc. Behav. Sci.* 98, 864–869. doi: 10.1016/j.sbspro.2014.03.493
- Kim, E. S., and Lee, E. J. (2021). A corpus-based analysis of lexical bundles between english L1 and L2 writers in medical journal abstracts: focusing on structures and functions. *J. Asia TEFL* 18, 142–160. doi: 10.18823/asiatefl.2021.18.1.9.142
- Kirsch, G. (1994). The politics of I-dropping. *College Compos. Commun.* 45, 381–383. doi: 10.2307/358817
- Kuo, C. H. (1999). The use of personal pronouns: role relationships in scientific journal articles. *Eng. Specific Purposes* 18, 121–138. doi: 10.1016/S0889-4906(97)00058-6
- Landis, J. R., and Koch, G. G. (1977). The measurement of observer agreement for categorical data. *Biometrics* 33, 159–174. doi: 10.2307/2529310
- Li, L., Franken, M., and Wu, S. (2018). Chinese postgraduates’ explanation of the sources of sentence initial bundles in their thesis writing. *RELC J.* 50, 37–52. doi: 10.1177/0033688217750641
- Li, L., Franken, M., and Wu, S. (2020). Bundle-driven move analysis: sentence initial lexical bundles in PhD abstracts. *Eng. Specific Purposes* 60, 85–97. doi: 10.1016/j.esp.2020.04.006
- Lin, P. (2010). “The phonology of formulaic sequences: a review,” in *Perspectives on Formulaic Language: Acquisition and Communication*, ed. D. Wood (London: Continuum), 174–194.
- Liu, C., and Chen, H. H. (2020). Analyzing the functions of lexical bundles in undergraduate academic lectures for pedagogical use. *Eng. Specific Purposes* 58, 122–137. doi: 10.1016/j.esp.2019.12.003

- Liu, J., and Lu, Y. (2019). A corpus-based comparative study on lexical bundles in native and Chinese scholars' English abstracts—taking linguistics and chemistry as an example. *Chin. J. Appl. Ling.* 42, 488–502. doi: 10.1515/CJAL-2019-0029
- Lou, B. (2010). Lexical bundles in learners' MA theses. *Contemp. For. Lang. Stud.* 9, 27–34.
- Lu, X., and Deng, J. (2019). With the rapid development: a contrastive analysis of lexical bundles in dissertation abstracts by Chinese and L1 English doctoral students. *J. Eng. Acad. Purposes* 39, 21–36. doi: 10.1016/j.jeap.2019.03.008
- Lyu, M., and Gee, R. W. (2020). Lexical bundles in thesis abstracts by L1 Chinese learners of English and U.S. students. *Eng. Lang. Teach.* 13, 141–155. doi: 10.5539/elt.v13n1p141
- Mauranen, A., Hynninen, N., and Ranta, E. (2010). English as an academic lingua franca: the ELFA project. *Eng. Specific Purposes* 29, 183–190. doi: 10.1016/j.esp.2009.10.001
- Myles, F. (2004). From data to theory: the over-representation of linguistic knowledge in SLA. *Trans. Philol. Soc.* 102, 139–168. doi: 10.1111/j.0079-1636.2004.00133.x
- Nekrasova, T. (2009). English L1 and L2 speakers' knowledge of lexical bundles. *Lang. Learn.* 59, 674–686. doi: 10.1111/j.1467-9922.2009.00520.x
- Niu, G. (2014). "Structurally and functionally comparative analysis of lexical bundles in the English abstracts of Chinese and international journals," in *Chinese Lexical Semantics*, eds X. Su and T. He (Switzerland: Springer), 349–357. doi: 10.1007/978-3-319-14331-6\_35
- Oakey, D. (2020). Phrases in EAP academic writing pedagogy: illuminating halliday's influence on research and practice. *J. Eng. Acad. Purposes* 44, 1–16. doi: 10.1016/j.jeap.2019.100829
- Omidian, T., Shahriari, H., and Siyanova-Chanturia, A. (2018). A cross-disciplinary investigation of multi-word expressions in the moves of research article abstracts. *J. Eng. Acad. Purposes* 36, 1–14. doi: 10.1016/j.jeap.2018.08.002
- Pan, F., Reppen, R., and Biber, D. (2016). Comparing patterns of L1 versus L2 English academic professionals: lexical bundles in telecommunications research journals. *J. Eng. Acad. Purposes* 21, 60–71. doi: 10.1016/j.jeap.2015.11.003
- Pan, F., Reppen, R., and Biber, D. (2020). Methodological issues in contrastive lexical bundle research: the influence of corpus design on bundle identification. *Int. J. Corpus Ling.* 25, 215–229. doi: 10.1075/ijcl.19063.pan
- Pang, P. (2009). A study on the use of four-word lexical bundles in argumentative essays by Chinese English-majors. *CELEA J.* 32, 25–45.
- Paquot, M., and Granger, S. (2012). Formulaic language in learner corpora. *Ann. Rev. Appl. Ling.* 32, 130–149. doi: 10.1017/S0267190512000098
- Pérez-Llantada, C. (2014). Formulaic language in L1 and L2 expert academic writing: convergent and divergent usage. *J. Eng. Acad. Purposes* 14, 84–94. doi: 10.1016/j.jeap.2014.01.002
- Qi, H., and Pan, F. (2020). Lexical bundle variation across moves in abstracts of medical research articles. *Southern Afri. Ling. Appl. Lang. Stud.* 38, 109–128. doi: 10.2989/16073614.2020.1763814
- Qin, J. (2014). Use of formulaic bundles by non-native English graduate writers and published authors in applied linguistics. *System* 42, 220–231. doi: 10.1016/j.system.2013.12.003
- Rayson, P. (2016). *Effect Size Calculator*. Lancaster, PA: Lancaster University.
- Rayson, P., and Garside, R. (2000). "Comparing corpora using frequency profiling," in *Proceedings of the Workshop on Comparing Corpora, Held in Conjunction With the 38th Annual Meeting of the Association for Computational Linguistics (ACL 2000)*, (Hong Kong), 1–6. doi: 10.3115/1117729.1117730
- Schmitt, N., and Carter, R. (2004). "Formulaic sequences in action: an introduction," in *Formulaic Sequences: Acquisition, Processing and Use*, ed. N. Schmitt (Amsterdam: John Benjamins), 1–22. doi: 10.1075/llt.9.02sch
- Scott, M. (2020). *WordSmith Tools: Version 8.0*. Oxford: Oxford University Press.
- Segalowitz, N. (2010). *Cognitive Bases of Second Language Fluency*. New York: Routledge. doi: 10.4324/9780203851357
- Simpson-Vlach, R., and Ellis, N. (2010). An academic formulas list: new methods in phraseology research. *Appl. Ling.* 31, 487–512. doi: 10.1093/applin/amp058
- Spencer, C. M., and Arbon, B. (1996). *Foundations of Writing: Developing Research and Academic Writing Skills*. Lincolnwood: National Textbook Co.
- Swales, J., and Feak, C. B. (2009). *Abstracts and the Writing of Abstracts*. Ann Arbor: The University of Michigan Press. doi: 10.3998/mpub.309332
- The University of Chicago Press Editorial Staff (2017). *The Chicago Manual of Style*, 17th Edn. Chicago: University of Chicago Press. doi: 10.7208/cmos17
- Titone, D., and Connine, C. (1999). On the compositional and noncompositional nature of idiomatic expressions. *J. Pragmat.* 31, 1655–1674. doi: 10.1016/S0378-2166(99)00008-9
- Wei, Y., and Lei, L. (2011). Lexical bundles in the academic writing of advanced Chinese EFL learners. *RELJ* 42, 155–166. doi: 10.1177/0033688211407295
- Wongwiwat, T. (2016). *Move Analysis and Lexical Bundle Analysis of Conference Abstracts: A Case Study of Thailand TESOL International Conference*. Ph.D. thesis. Phra Nakhon: Thammasat University.
- Wray, A. (2002). *Formulaic Language and the Lexicon*. New York: Cambridge University Press. doi: 10.1017/CBO9780511519772
- Xu, F. (2012). The use and developmental features of lexical bundles in Chinese learners' English academic writing. *For. Lang. China* 9, 51–56.

**Conflict of Interest:** The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

**Publisher's Note:** All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Copyright © 2022 Bao and Liu. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.



# Integrating Automatic Speech Recognition Technology Into Vocabulary Learning in a Flipped English Class for Chinese College Students

Michael Yi-Chao Jiang<sup>1,2</sup>, Morris Siu-Yung Jong<sup>1,2</sup>, Na Wu<sup>3</sup>, Bin Shen<sup>4,5\*</sup>, Ching-Sing Chai<sup>1</sup>, Wilfred Wing-Fat Lau<sup>1</sup> and Biyun Huang<sup>2</sup>

<sup>1</sup> Department of Curriculum and Instruction, Faculty of Education, The Chinese University of Hong Kong, Hong Kong, China, <sup>2</sup> Centre for Learning Sciences and Technologies, Hong Kong Institute of Educational Research, The Chinese University of Hong Kong, Hong Kong, China, <sup>3</sup> College of International Business, Shenyang Normal University, Shenyang, China, <sup>4</sup> School of Foreign Languages, Fuzhou University, Fuzhou, China, <sup>5</sup> Center for Foreign Language Education and Teaching, School of Foreign Languages, Fuzhou University, Fuzhou, China

## OPEN ACCESS

### Edited by:

Ehsan Rassaei,  
Majan University College, Oman

### Reviewed by:

Ali Soyoof,  
Monash University, Australia  
Mostafa Azari Noughabi,  
University of Gonabad, Iran  
Habib Soleimani,  
University of Kurdistan, Iran

### \*Correspondence:

Bin Shen  
80176713@qq.com;  
bin.shen@fzu.edu.cn

### Specialty section:

This article was submitted to  
Language Sciences,  
a section of the journal  
Frontiers in Psychology

**Received:** 23 March 2022

**Accepted:** 06 June 2022

**Published:** 15 July 2022

### Citation:

Jiang MY-C, Jong MS-Y, Wu N, Shen B, Chai C-S, Lau WW-F and Huang B (2022) Integrating Automatic Speech Recognition Technology Into Vocabulary Learning in a Flipped English Class for Chinese College Students. *Front. Psychol.* 13:902429. doi: 10.3389/fpsyg.2022.902429

Although the automatic speech recognition (ASR) technology is increasingly used for commercial purposes, its impact on language learning has not been extensively studied. Underpinned by the sociocultural theory, the present work examined the effects of leveraging ASR technology to support English vocabulary learning in a tertiary flipped setting. A control group and an experimental group of college students participated in a 14-week study. Both groups had their English classes in a flipped fashion, but the experimental group was assigned with ASR-assisted oral tasks for pre-class self-learning. The pre- and post-intervention in-class task performance of both groups was audio-recorded and transcribed for data analysis. The triadic complexity-accuracy-fluency (CAF) framework was adopted to evaluate the participants' vocabulary learning. The between- and within-subjects effects were examined mainly through procedures of MANCOVA and mixed-design repeated measures ANCOVA. Results showed that on all the metrics of lexical complexity and speed fluency, the experimental group outperformed the control group, and had significant growth over time. On the other hand, the control group only improved significantly overtime on the G-index. On lexical accuracy, there was no significant difference between the two groups, and the within-subjects effect was not significant for either group. The findings lent some support to Skehan's Trade-off Hypothesis and discussions were conducted regarding the triarchic CAF framework.

**Keywords:** automatic speech recognition, flipped classroom, vocabulary learning, CAF framework, trade-off effect

## INTRODUCTION

Learning a second or foreign language (L2/FL) usually requires a substantial amount of constant corrective feedback from a source other than learners' perceptions (Franco et al., 2010). Given that sounds are filtered through their mother tongue (McCrocklin, 2016), L2/FL learners are quite unlikely to monitor their own oral speech practice in the target language. In addition, the



feedback provided by language teachers for each learner is oftentimes subject to time and space restraints. With the advancement of educational technology and artificial intelligence, especially in the domain of intelligent computer-assisted language learning (iCALL), the automatic speech recognition (ASR) technology is progressively regarded as a conceivable solution to address that issue (Mroz, 2018; McCrocklin, 2019; Evers and Chen, 2020; Jiang et al., 2021; Zhai et al., 2021). ASR-based technologies and applications attract L2/FL researchers' and practitioners' attention (McCrocklin, 2016; Penning de Vries et al., 2020), owing to the features such as considerable amounts of practice, consistent and unbiased feedback, and diverse forms of visual representations (Levis, 2007). According to Rassaei (2021), those features are integrally linked with and emphasized in sociocultural theory as critical characteristics of an effective classroom teaching. Moreover, apart from more opportunities for extensive interaction in the target language and real-time feedback, ASR-based technology can also provide L2/FL learners with more control over their self-learning, thus creating a less threatening self-paced environment for individual learners when learning to speak in the target language (Jiang et al., 2021).

A growing number of empirical studies examining the effect of ASR technology on L2/FL learning have been conducted in the past decade with the majority of them dedicated to research on L2/FL pronunciation (McCrocklin, 2016; Evers and Chen, 2020), some dedicated to improving oral grammatical skills and complexity (Penning de Vries et al., 2020; Jiang et al., 2021), while far less to vocabulary learning (Bashori et al., 2021), despite the paramount role of productive vocabulary learning in L2/FL learning (Schmitt, 2010; Li and Hafner, 2021). Previous research has also lent some support to the effects of iCALL technologies on vocabulary knowledge (e.g., Chen and Hsu, 2019; Soyoof et al., 2022), thus making the present attempt to integrate ASR technology into vocabulary learning in an L2/FL classroom deserve attention. Moreover, most of the studies were carried out in conventional L2/FL settings, where the use of the ASR-based technology might not be to the fullest due to the limited in-class time or less-structured pre-class self-learning (Jiang et al., 2020). Besides, those studies mainly focused on the technological integration of the ASR-based applications, but few detailed the instructional design and implementation as much. Because of such insufficient description of how the ASR-based technology was pedagogically integrated into task-based language learning in those studies, their pedagogical implications for future studies may be diminished. Furthermore, an overwhelming majority of the studies utilized self-reported data (e.g., attitudes toward the use of ASR technology) and adopted only overall measures of the students' academic performance (e.g., overall assessment of oral proficiency) to examine the effects of ASR technology. In contrast, objective and fine-grained measures such as those based on the tripartite framework of complexity, accuracy and fluency (CAF) (Skehan, 1996) are barely used. Specific measures of learners' linguistic performance are regarded as more direct gauges in depicting students' language learning, and thus may further contribute to diversifying the instructional design of the flipped classroom approach (Jiang et al., 2021). Therefore,

little is known about how L2/FL learners' vocabulary learning is affected by the ASR-based technology. With those research gaps identified, the present study aimed to utilize the CAF framework to investigate how the ASR-based technology might facilitate students' L2/FL vocabulary learning. The findings may contribute to a better understanding of exploiting ASR-based applications and shed light on the course design in flipped EFL classrooms.

## RELATED WORKS

### A Sociocultural Theoretical (SCT) Perspective

In the past decades, there emerges an accumulating interest in taking a sociocultural theoretical perspective to research L2/FL learning (Rassaei, 2014, 2020). According to the sociocultural approach, language development is rooted in dialogic interactions (Ellis, 2009), and learners are empowered to perform challenging tasks which may exceed their abilities through social interaction with assistance from other capable learners or social environment and artifacts (Vygotsky, 1978). Traditionally associated with Vygotsky's (1978, 1987) work, SCT relates social interaction to individual cognitive development. Central to the notion of SCT lies the stance that higher forms of learning and thinking originate from social interaction (Vygotsky, 1978; Villamil and de Guerrero, 2006). Different from other second language acquisition (SLA) and cognitive theories such as the information-processing approach and interactionist theory, which view social interaction and information processing from such interaction as separate practices though admitting the significant role of social interaction in second language learning, SCT holds "social interaction (with both humans and artifacts participating dialogically) is internalized, the external-dialogic becomes the internal-dialogic, and a socially constructed dialogic mind emerges" (Villamil and de Guerrero, 2006, p. 24). Vygotsky's (1978) zone of proximal development (ZPD) is defined as the distance between what a learner can do with assistance and what the same learner can perform independently. Put it another way, there are thereby distinctions between a learner's actual level of language learning improvement when engaged in self-learning without external support and his/her potential level of development when facilitated by assisted and collaborative performance. From this perspective, the instant feedback on language production as generated by the ASR software could be conceptualized as the social artifact/mediator, interaction with which can lead to growth and improvement in vocabulary learning on the learners' side.

Important SCT concepts to understand and investigate the potential effects of ASR-enhanced technology on vocabulary learning in an L2/FL context include mediation, internalization, and developmental change. As Villamil and de Guerrero (2006) analyzed, the adult human mind has to firstly go through a sociocultural mediation to transform from lower forms of thinking (natural memory, basic perception) to higher forms of thinking (logical reasoning, problem solving). Mediation by others, mediation by self, and mediation by artifacts are the three forms of mediation postulated by SCT;

moreover, internalization of mediation is a developmental process to achieve higher order of thinking, and Wertsch's (1979) categorization of regulation stages proposed learners move from other-regulation to self-regulation in the transition from interpsychological to intrapsychological activity. To be specific, when ASR technology is integrated into vocabulary learning, learners could receive assistance from social artifact (feedback from ASR software) and go through a sociocultural mediation by artifacts and self to transform from lower to higher forms of thinking, move from other-regulation (i.e., performing with assistance from ASR software) to self-regulation (i.e., capable of independent performance of oral task) in the transition from interpsychological to intrapsychological activity.

## ASR-Assisted Vocabulary Learning in Flipped EFL Classrooms

Owing to their easy accessibility and ubiquity, smartphones and tablets can be utilized for providing constant feedback and mediation to language learners (Rassaei, 2021), and the past decades have witnessed a research boom in the field of iCALL. In particular, ASR has emerged as one of the more promising iCALL technologies which is empowered by computer-based processes of decoding and transcribing oral language usually into text form (Kim, 2006). When ASR technology is integrated in a pedagogically sound way, it facilitates interactive learning environments (Wang and Young, 2014), offers instant assessment and feedback on language pronunciation and language use (Franco et al., 2010), enables easily accessible oral practice opportunities beyond time and space limitations (Torlakovic and Deugo, 2004), and reduces L2/FL speaking anxiety (Bashori et al., 2020). With these merits noted, ASR is considered beneficial for L2/FL oral practice. Specifically, mounting evidence has been accumulated on the effectiveness of applying ASR to enhance L2/FL pronunciation (Neri et al., 2008; McCrocklin, 2016) and new and sporadic attempts were made on improving oral grammatical skills (Penning de Vries et al., 2014, 2020). However, notably scant attention has been focused on the employment of ASR in promoting vocabulary learning in the L2/FL learning contexts (Bashori et al., 2021).

In view of the tenet of the flipped classroom approach and the Bloom's taxonomy (Anderson and Krathwohl, 2001), lower-level learning objectives (i.e., remembering- and understanding-oriented content), such as vocabulary learning in L2/FL classrooms, could be achieved through students' pre-class self-study. Flipped classroom is considered a well-matched approach for incorporating ASR-based technology in L2/FL vocabulary learning because the pre-class self-learning is usually well-organized and more self-paced than in a conventional classroom. Moreover, students in a flipped classroom are expected to spend adequate time in self-learning and practicing prior to attending class (Jong, 2017; Jong et al., 2019). Consequently, empirical studies are needed to examine how flipped classroom approach could facilitate the integration of ASR technology in the context of L2/FL learning.

Vocabulary plays a critical role in L2/FL learning given vocabulary knowledge being found to significantly predict the four essential language skills (Schmitt, 2010; Milton, 2013). But for learners of English as a second or foreign language (ESL/EFL), vocabulary acquisition often poses a challenging burden (Lo and Murphy, 2010; Webb and Nation, 2017). Moreover, L2/FL class time usually appears inadequate for vocabulary learning (Nation, 2006). In most cases, L2/FL learners may need to seek alternative resources to learn words independently out of class (Teng, 2020). Luckily, technological advancements have induced such learning opportunities with iCALL approaches, such as captioned videos (Teng, 2019, 2022), mobile games (Chen and Hsu, 2019; Abdulrahman and Jullian, 2020; Rahman and Angraeni, 2020) and virtual reality tools (Madini and Alshaikhi, 2017; Tai et al., 2020). Results have attested to the positive effects of technologies on learners' vocabulary knowledge, especially productive vocabulary learning, and their self-efficacy in vocabulary learning (e.g., Li and Hafner, 2021; Soleimani et al., 2022), which could pave the way for integrating ASR into vocabulary learning. Among the scarce endeavors, Bashori et al. (2021) conducted a quasi-experimental study with Indonesian secondary school students and reported students from the two ASR intervention groups (using two different ASR websites) outperformed the regular class group in their knowledge of the targeted vocabulary and emotional states (i.e., anxiety and enjoyment).

It is worthwhile to note in Bashori et al.'s (2021) study, vocabulary knowledge was assessed using written vocabulary test on the targeted words. This approach presented a relatively simplified assessment of the learners' mastery of the targeted vocabulary in an arbitrary fashion of correct or incorrect answers while jeopardizing an informative insight into the multi-dimensional construct of vocabulary competence. With the endorsement of the multi-componential nature of linguistic competence (Norris and Ortega, 2009), it is important for researchers to examine domain-specific outcome measures, including L2 complexity, accuracy and fluency. For example, learners' lexical and syntactic complexity in English oral performance was noticeably improved when engaged in ASR-based oral tasks for a semester (Jiang et al., 2021). Feedback generated from iCALL-based speaking practices could lead to more accurate utterances (Mackey and Goo, 2007). ASR-based pronunciation system was found to be equally capable of diagnosing human pronunciation errors as human raters did at the segmental level, and it was found that learners' varied pronunciation learning needs were met by using the ASR technology (Xiao and Park, 2021). That said, to bridge the research void, we intend to adopt the triadic CAF framework to measure the learners' vocabulary development, in the hope of retrieving a detailed diagnostic evaluation of their vocabulary learning as a result of ASR-enhanced oral practice.

## The Triadic Componential CAF Framework

Language proficiency is perceived as a multi-componential, multilayered, and multifaceted construct rather than a unitary one, and its principal components can be fruitfully captured by the framing of CAF (Housen et al., 2012; Jiang et al., 2021).

Skehan (1996, 1998) theoretically combined the three constructs into one proficiency model and provided the working definitions which are still in use in areas such as SLA. Ever since, a heated debate has surrounded the issue of quantifying language learners' output in both written and spoken form.

Complexity concerns size, elaborateness, richness, and diversity of the language learners' linguistic system (Bui, 2021). In the literature pertaining to CAF, complexity is generally assessed through the competence to use a wide and varied range of advanced vocabulary and sophisticated structures in the target language (Skehan, 1998; Ellis, 2003, 2008; Housen et al., 2012). Because of its polysemous nature, complexity in language learning retains multiple meanings (Michel, 2017; Bulté and Roothoof, 2020) and is the most debated construct of the CAF triad (Pallotti, 2009). Following Michel (2017), complexity can be applied to three different dimensions, i.e., developmental, cognitive and linguistic complexity. Empirical studies converge to show that linguistic complexity is the most commonly measurable construct (Bui, 2021). Operationally, measures created for assessing linguistic complexity are dichotomized into two broad categories: lexical complexity and syntactic complexity. In the context of this study, lexical complexity is adopted as an indicator of EFL learners' vocabulary learning performance. In literature, a considerable number of EFL studies have investigated the role of lexical complexity in language learning, but most of their data were written English (e.g., Barrot and Gabinete, 2021; Han et al., 2021). In contrast, few studies have investigated oral lexical complexity in the context of EFL learning (Bulté and Roothoof, 2020). In response, the present study seeks to address this gap by examining EFL learners' lexical complexity in their oral English. Although lexical complexity can be investigated through various aspects such as diversity, density and sophistication (Skehan, 2003; Bulté and Housen, 2012), complexity is primarily shaped by lexical diversity, and operationally, lexical diversity is usually the most frequently used measure for assessing lexical complexity (e.g., Ågren et al., 2012; Han et al., 2021) and lexical development (e.g., Crossley et al., 2009). In the current study, therefore, as has been the case for most studies so far, complexity is quantified using lexical diversity in the tripartite CAF framework.

Accuracy is arguably the most straightforward and internally consistent construct of the CAF triad (Housen and Kuiken, 2009). Fundamentally, accuracy is generally defined as the degree to which a learner's language performance (and the target language system that underlies his or her performance) deviates from the native-like use (Wolfe-Quintero et al., 1998; Pallotti, 2009; Barrot and Gabinete, 2021). The deviations are typically labeled as "errors" and based on the classification of the errors, the measures of linguistic accuracy may concern lexis, morphology, syntax, phonology, and pragmatics (Chavez, 2014). In L2/FL studies, researchers showcased that accuracy could be reliably and validly measured by error-free metrics, such as calculating the number of error-free clauses of all clauses or the ratio of error-free T-units to all T-units (Polio and Shea, 2014; Barrot and Gabinete, 2021). But in China, English is learned and used as an FL rather than an L2, and for most non-English major undergraduates, the proportion of

error-free utterances generated in their classroom talk is assumed to be exceedingly low. Therefore, given the potential floor effect of error-free metrics, the present study adopted error-based metrics for quantifying accuracy. According to previous studies (e.g., Liao, 2020), lexical accuracy and morphosyntactic accuracy are two most frequently adopted metrics. To be specific, lexical accuracy involves an ability to retrieve an appropriate word and use it correctly in a specific context, while morphosyntactic accuracy focuses on aspects such as agreement, inflection as well as retrieving an appropriate structure or organizing constituents in order. In the context of the present study, therefore, lexical accuracy is adopted as an accuracy indicator to measure students' vocabulary language learning, which was operationalized through students' lexical errors in their oral English.

In general usage, fluency is often understood as a language learner's overall language proficiency that particularly relates to the ease, eloquence, smoothness and native-likeness of their speech or writing (Lennon, 1990; Chambers, 1997; Van Waes and Mariëlle, 2015). A fluent L2/FL learner is believed to be capable of producing the target language with native-like rapidity, pausing, hesitation or reformulation. In contrast to complexity and accuracy, which are primarily associated with the current state of the learner's interlanguage knowledge, fluency is oftentimes a phonological phenomenon (Housen et al., 2012). Likewise, fluency is also multi-dimensional as the other two constructs in the tripartite CAF framework. Following Skehan and other researchers (Skehan, 2003, 2009; Tavakoli, 2016; Tavakoli et al., 2016), fluency can be examined through its subdimensions such as speed fluency (the rate and density of linguistic units produced), breakdown fluency (number, length, and location of pauses) and repair fluency (false starts, misformulations, self-corrections, and repetitions) (Housen et al., 2012). According to Lambert and Kormos (2014), fluency metrics that are conceptualized based on speech rate (i.e., a ratio of syllables produced to time taken to produce them) are the most frequently used measures. Conversely, dysfluency metrics (i.e., breakdown fluency, repair fluency) that are based on filled/unfilled pauses, hesitations, false starts, and so on did not show a strong association with learners' overall oral proficiency assessed by native speakers (Kormos and Dénes, 2004). Moreover, compared with breakdown or repair fluency, speed fluency is more linked to the L2/FL lexicon in oral output, and therefore it is posited to be a manifestation of a more advanced proficiency level pertaining to vocabulary learning. As such, in the context of the present study, speed fluency was employed as a fluency indicator to evaluate students' vocabulary learning performance.

## The Trade-Off Hypothesis in Task-Based Language Learning

Earlier known as the Limited Attentional Capacity Model, the Trade-off Hypothesis (Skehan, 2009) states that learners' attentional resources are limited, and interlocutors must allot their attentional resources a task requires during the processes (Sample and Michel, 2014). As a result, if task demands



exceed the available attentional resources, learners' linguistic performance in terms of complexity, accuracy, and fluency may compete with each other (Sample and Michel, 2014; Sun and Révész, 2021). Particularly, it has been argued that a trade-off exists between attention to form and attention to meaning during task performance (Skehan, 1998, 2009; Skehan and Foster, 2001).

For communicative purposes, L2/FL learners are assumed to prioritize meaning (i.e., fluency) over form (i.e., accuracy and complexity) (Skehan, 2009). In other words, performing L2/FL tasks may lead to conflicts between meaning and form for learners' attentional resources. Therefore, when learners concentrate on being fluent in delivering the communicative content, fewer attentional resources will be available for producing complex and accurate utterances (Sample and Michel, 2014). Furthermore, following Skehan (2009), a further trade-off is likely to arise between these two latter dimensions because learners may lack resources to pay attention to both complexity and accuracy simultaneously.

As aforementioned, the incorporation of ASR-technology provides L2/FL learners with an avenue of repeated practice based on synchronic feedback, preparing themselves in advance for the in-class higher-order tasks. To the best of our knowledge, however, few studies have been conducted to investigate how the ASR technology influences EFL learners' linguistic performance, especially in the domain of vocabulary learning. It also remains unclear whether the trade-off effect still holds when ASR-based technology is incorporated into task-based language learning. The current study, therefore, aims to fill this gap by testing the Trade-off Hypothesis in the context of ASR-enhanced task-based language learning with a research focus on vocabulary learning. Based on the research gaps identified, three research questions (RQ) were formulated in the current study:

RQ 1: Does the ASR-based technology embedded in pre-class self-study lead to differences in EFL learners' lexical complexity in a flipped classroom?

RQ 2: Does the ASR-based technology embedded in pre-class self-study lead to differences in EFL learners' lexical accuracy in a flipped classroom?

RQ 3: Does the ASR-based technology embedded in pre-class self-study lead to differences in EFL learners' speed fluency in a flipped classroom?

## METHODS

### Participants

Sixty-three first-year undergraduates of two EFL classes in a Chinese university were recruited in the quasi-experiment. Their majors included Chinese literature and arts, sociology, public administration and management, education, computer science and technology, biological engineering, law and mathematics. Their average age was 18.1 years old; 17.5% of them were male, and 82.5% were female. All the students consented to participate in the study approved by the research site university. According to the pre-intervention survey, the participants had English learning history for approximately 11 years on average and they reported an average score of 128.5 (out of 150) for

their college entrance English examination, indicating that on a general basis they were ready to learn English at the tertiary level. With regard to their experiences of flipped learning, 90.5% of the students had "never" or "seldom" learned in a flipped fashion, and 9.5% "some" experience of flipped learning. Moreover, 65.1% of the participants reported "no" or "little" training specific for oral English, and 27% "some" experiences of oral English learning; only 7.9% had "sufficient" training in oral English back at high school.

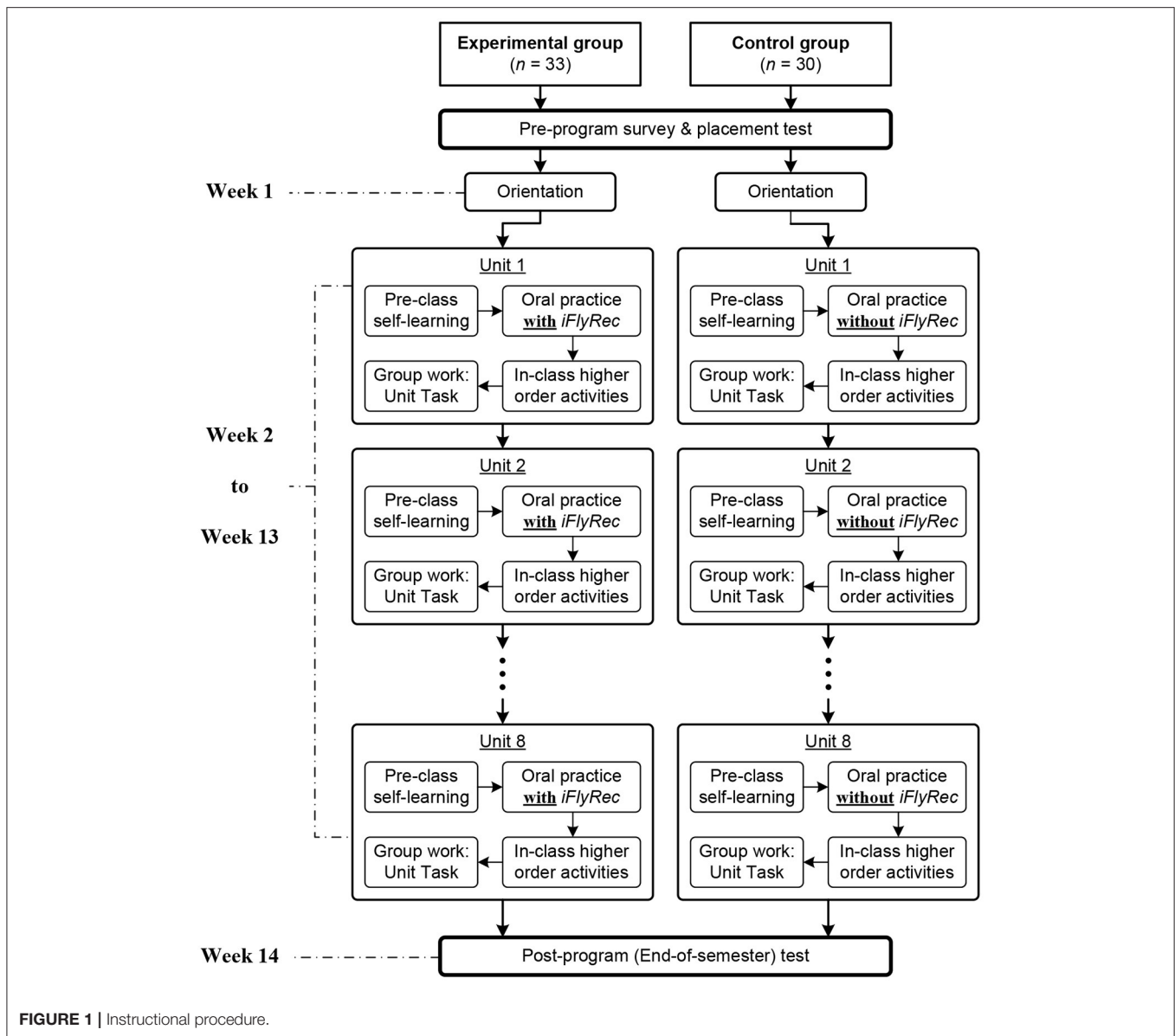
### Course Design

This course was part of the College English program for Year 1 and Year 2 undergraduates, which aimed to develop learners' English proficiency and foster their English skills for both general and academic purposes. Each semester, the course covered a total of eight learning units. An online learning platform, i.e., Unipus (<https://u.unipus.cn/>), developed by the course book publisher, was utilized for the flipped implementation for both classes. All the course contents (i.e., vocabulary, cultural background information, texts and recordings, in-class tasks and post-class assignments) on Unipus were accessible with smart devices such as smartphones or tablets. On a weekly basis, the students in both classes had a 90-min face-to-face session with the same EFL teacher who had been teaching the program for ten consecutive years. Within each class, the students were randomly assigned into workgroups of three or four for performing group-based tasks, and for data collection reasons, the composition of the workgroups remained unchanged until the end of the semester.

Each learning unit consisted of several sections with varied learning tasks. According to Bloom's taxonomy (Anderson and Krathwohl, 2001), some of the sections and tasks were understanding- and remembering-oriented, such as *Reading Across Cultures* and *Language in Use*, while others were more applying-, analyzing-, evaluating-, and creating-oriented, such as *Reading Skills*, *Guided Writing* and *Unit Task (UT)*. In light of the rationale of flipped classroom approach, tasks that were at the lower level of the taxonomy (understanding- and remembering-oriented) were flipped outside the classroom for students' pre-class self-learning on Unipus. Conversely, tasks at the higher level (e.g., applying- and analyzing-oriented activities) were performed in class (Jong, 2019a; Jong et al., 2022). In particular, a comprehensive UT was performed in class toward the end of each learning unit. It was a production-oriented group activity for the students to conduct a topic-based discussion that required higher-order language skills such as analyzing and evaluating. The performance of each group member and their peer interaction while performing the UT were audio recorded as the major data source in this study. The instructional procedure is demonstrated in **Figure 1**.

An ASR-based application called iFlyRec (<https://www.iflyrec.com>), which is developed by iFlyTek, a well-known intelligent speech and artificial intelligence company, was utilized in the study. iFlyRec is free to download and can run on iOS and Android systems. One of its salient features lies in its real-time conversion of speech into text in multiple languages and even some Chinese dialects (**Figure 2**). Moreover, it also supports interlingual translation in oral form in several languages





including Chinese, English, Korean, Japanese, and Russian. In the current study, the students in the experimental group were required to perform oral tasks in pre-class self-learning with the assistance of iFlyRec. Based on the immediately transcribed texts as feedback for their oral speech, the experimental group students were encouraged to repeat their practice until their utterance was fully understood by the application. Contrarily, the students in the control group performed the same sets of pre-class oral tasks with no ASR-based applications. They needed to evaluate their oral performance by themselves. When they felt that their oral tasks were satisfactorily performed, the students in both groups should upload their recordings of their completed tasks to Unipus for assessment before the next class. Enlightened by the sociocultural theory, the integration of the ASR-based practices into students' pre-class self-learning is intended to promote their vocabulary learning, based on the ZPD assumption

that there are gaps between learners' original level of language learning and their potential level of learning development when facilitated by social interaction with artifact/mediation. It is therefore hypothesized that with the instant feedback on language production afforded, the ASR technology could provide opportunities for mediated performance that can make a difference in vocabulary learning on the learners' side.

## Research Design

This study adopted a pre-and-post quasi-experimental design. The independent variable was the group factor of two levels, and the dependent variable was the participants' linguistic performance in relation to vocabulary learning (coded from their UT performance). The two classes were randomly determined as the experimental group ( $n = 33$ ) and the control group ( $n = 30$ ). A survey was administered before the experiment to gather

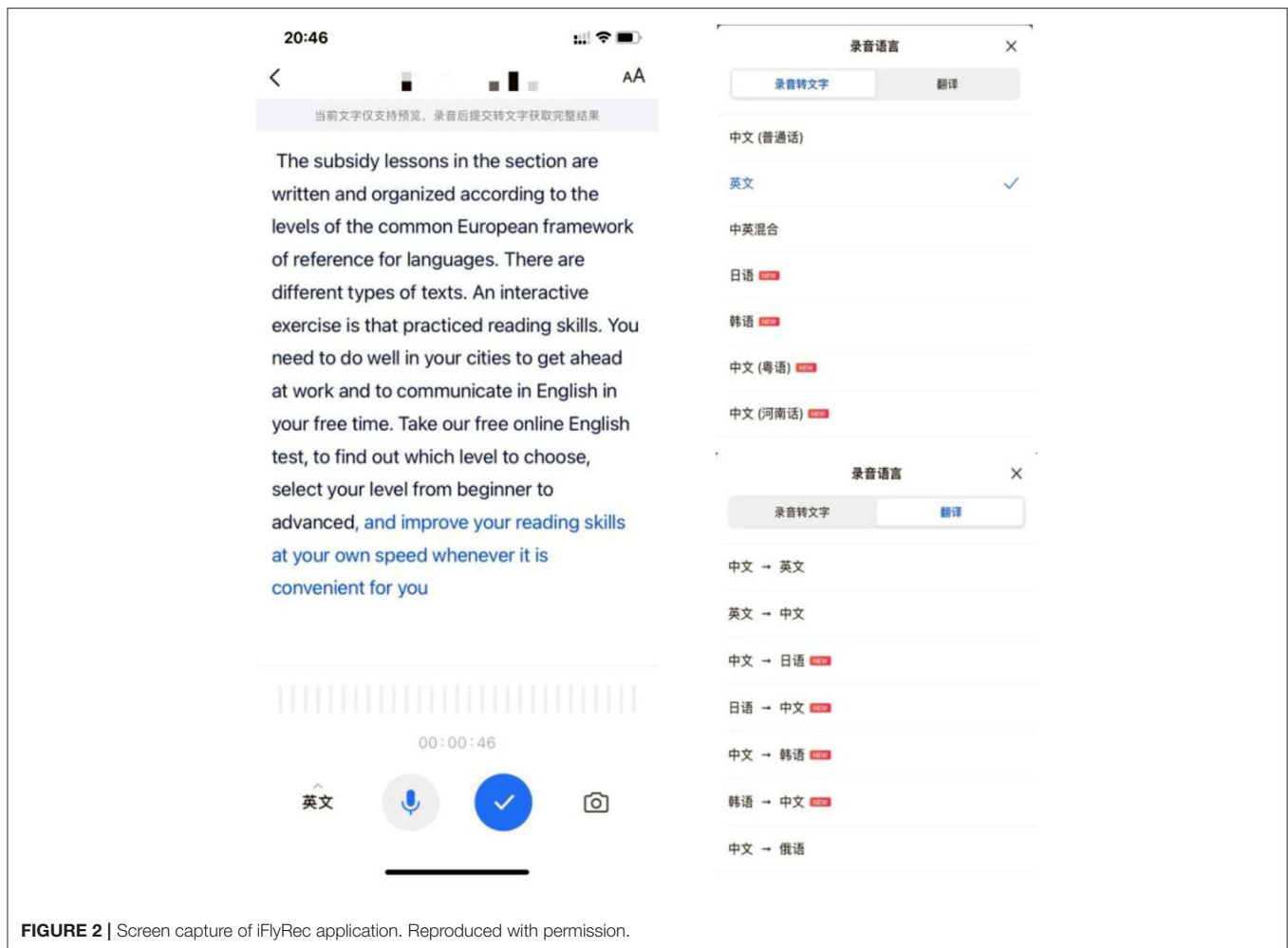


FIGURE 2 | Screen capture of iFlyRec application. Reproduced with permission.

the participants' background information and a placement test was used to measure their pre-intervention English proficiency, which was controlled for as a covariate in the data analysis. Accordingly, MANCOVA and mixed-design repeated measures ANCOVA procedures were adopted in this study to examine the between- and within-subjects effects.

Given that the participants might not know what was expected of them in a college EFL classroom when they just started college learning, their UT performance in Unit 1 was not collected in the study. Instead, their UT performance of Unit 2 was used as the pre-intervention data, and their performance of Unit 8 (i.e., the last unit of the semester) as the post-intervention data. The task performance of each workgroup was audio recorded while the students were performing the UTs. The recordings were transcribed into searchable text form and then coded with ELAN (<https://tla.mpi.nl/tools/tla-tools/elan>), a professional annotation tool for audio and video recordings. In data preprocessing, seven participants (four from the experimental group and three from the control group) were excluded because of their recording quality, dropout of the program or absence in class. Consequently, pre- and post-intervention recordings of 56 students (29 from the experimental group and 27 from

the control group) were ultimately transcribed and coded for further analysis. The students were invited to proofread the transcriptions of their recordings to ensure the accuracy of the transcribed texts.

## Measures and Instruments

Based on the CAF framework, the participants' vocabulary learning performance was operationalized as lexical diversity, lexical accuracy and speed fluency (Table 1) in the current study. Specifically, lexical diversity was assessed through both simple metric (i.e., G-index) and complex metrics (i.e., vocd-D and MTLD), respectively. Lexical accuracy was quantified by the number of lexical errors against the analysis of speech unit (AS-unit), and speed fluency was estimated by unpruned speech rate, i.e., syllables per minute including all the utterances.

G-index, a widely utilized simple indicator of lexical complexity is obtained by dividing the types (the total number of different words) occurring in a speech or text sample by the square root of its tokens (the total number of words) (Guiraud, 1960). However, quantitative linguistic studies have shown that measures based on type/token ratio (TTR) are flawed and subject to the length of the text sample (see Richards and Malvern, 1997;

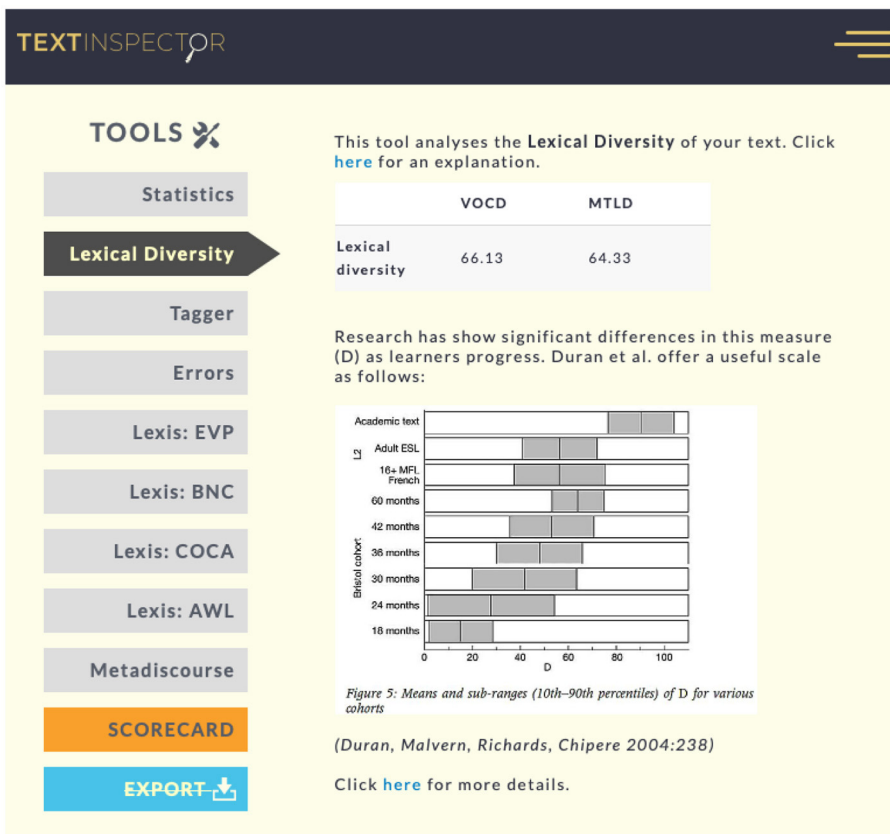
Tweedie and Baayen, 1998, for a demonstration). In response, we employed two more complex metrics, i.e., vocd-D and the measure of textual lexical diversity (MTLD), which are based on mathematical probabilistic models and are not susceptible to text length. They are calculated through computer programs and are results of a series of random text samplings. The vocd-D value has been used in numerous studies, although deBoer (2014) cautioned that vocd-D was still affected by text length and might be less reliable outside of an ideal range of perhaps 100–500 words. Conversely, McCarthy and Jarvis (2007) demonstrated that MTLD was a powerful index of lexical diversity, but further research was needed to confirm their findings in a range of settings. As such, the present study adopted both metrics in the hope of obtaining a clearer picture of the data and avoiding

drawing false conclusions. As was suggested by McCarthy and Jarvis (2010), researchers should use these indices together rather than any single index because lexical complexity can be assessed in many ways, and each approach may be informative as to the construct under investigation. Besides, the calculation of these two metrics involved all the words produced by the interlocutors (Albert, 2011). TextInspector (<https://textinspector.com>), a professional online tool for analyzing linguistic data, was utilized to calculate vocd-D and MTLD (Figure 3). It also provided basic statistics of a given text, such as TTR, syllable count, average sentence length and so on.

In terms of operationalizing lexical accuracy in this study, AS-unit was employed as the production unit, referring to “a single speaker’s utterance consisting of an independent clause, or sub-clausal unit, together with any subordinate clause(s) associated with either” (Foster et al., 2000). It is a length-based production unit specifically proposed as an improved alternative for oral discourse segmentation in SLA (Norris and Ortega, 2009; Jiang et al., 2021). Compared with other production units in use (e.g., C-unit, T-unit; see Foster et al., 2000 for details), the AS-unit is adequate and reliable when applied to transcriptions of complex oral data, which tend not to lend themselves easily to a clear division into units (Foster et al., 2000), especially for non-native speakers of English. The lexical errors (e.g., retrieve

**TABLE 1** | Metrics for measuring vocabulary learning performance.

CAF components	Sub-dimensions	Metrics
Complexity	Lexical diversity	G-index, vocd-D, MTLD
Accuracy	Lexical accuracy	Lexical errors per AS-unit
Fluency	Speed fluency	Unpruned syllables articulated per minute



**FIGURE 3** | Screen capture of TextInspector. Reproduced with permission.

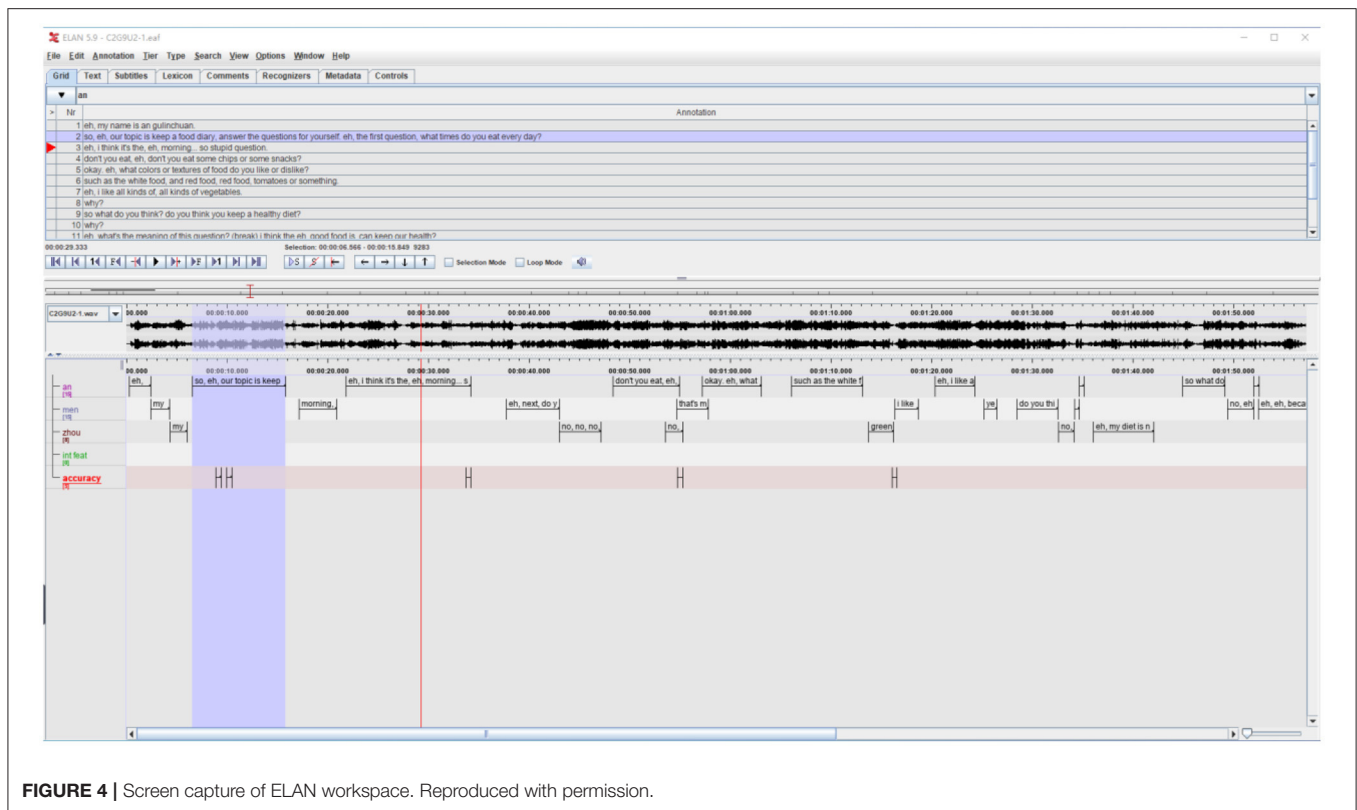


FIGURE 4 | Screen capture of ELAN workspace. Reproduced with permission.

inappropriate words or use them incorrectly in a specific context) were coded with ELAN, a piece of professional software for annotating audio and video recordings (Figure 4). One author and the course teacher conducted the coding and cross-checked the results. Any disagreement between the two coders was resolved through discussion until a consensus was reached. The current study was part of a doctoral study that involved more CAF metrics, and the overall inter-rater reliability was estimated through Krippendorff's  $\alpha$  (Hayes and Krippendorff, 2007) and was computed to be 0.818 ( $>0.8$ ), indicating consistency between the two coders.

The speed fluency was estimated by unpruned speech rate (i.e., syllables per minute including all the utterances such as false starts, self-corrections, and repetitions), which was computed by dividing the number of all the syllables produced by the time taken to produce them. As aforementioned, the count of syllables was reported as a basic statistic by TextInspector, and the time taken for each interlocutor could be easily read through ELAN after annotating the audio clips.

## RESULTS

### Lexical Complexity

Lexical complexity was estimated through G-index, vocd-D value and MTLD, of which the descriptive results were tabulated below (Table 2). MANCOVA was performed to examine the between-subjects effects with the pre-intervention placement test score as a covariate. Results showed that at the significance level of 0.05

(\* indicates  $p < 0.05$ , \*\*  $p < 0.01$ , \*\*\*  $p < 0.001$ ), the two groups had no significant differences in any of the three metrics of lexical diversity when performing their first UT. Conversely, after the intervention of a semester, the students in the experimental group significantly outscored their counterparts in the control group on G-index ( $F = 6.571^*$ ;  $p = 0.013 < 0.05$ ), vocd-D ( $F = 12.502^{***}$ ;  $p < 0.001$ ), and MTLD ( $F = 4.627^*$ ;  $p = 0.036 < 0.05$ ) when performing the last UT. The corresponding effect sizes (estimated by partial  $\eta^2$ ) were calculated to be 0.110 for G-index, 0.191 for vocd-D and 0.080 for MTLD, respectively, which all indicated medium to large effect sizes of the intervention on students' lexical diversity. Following Cohen (1988) and Miles and Shevlin (2001), the thresholds of partial  $\eta^2$  adopted in this study are small partial  $\eta^2 > 0.01$ , medium  $> 0.06$ , and large  $> 0.14$ .

Mixed-design repeated measures ANCOVAs were performed to further examine the corresponding between- and within-subjects effects. The results showed a significant between-subjects effect on vocd-D ( $F = 5.744^*$ ;  $p = 0.020 < 0.05$ ) and MTLD ( $F = 4.293^*$ ;  $p = 0.043 < 0.05$ ). Conversely, no significant between-subjects effect was noticed on G-index ( $F = 2.691^*$ ;  $p = 0.107 > 0.05$ ). Follow-up simple-effect tests revealed that the experimental group had a significant improvement on G-index ( $t = 7.994^{***}$ ,  $p < 0.001$ ) and MTLD ( $t = 2.271^*$ ,  $p = 0.031 < 0.05$ ) and a marginally significant improvement on vocd-D ( $t = 1.914$ ,  $p = 0.066 < 0.1$ ). Conversely, in the control group, only the improvement on G-index was statistically significant ( $t = 5.051^{***}$ ,  $p < 0.001$ ); no statistically significant change was witnessed on



vocd-D ( $t = 0.944$ ,  $p = 0.354 > 0.05$ ) or MTLD ( $t = 1.345$ ,  $p = 0.190 > 0.05$ ) (Figure 5).

## Lexical Accuracy

Lexical accuracy was quantified through an error-based metric, i.e., the number of lexical errors per AS-unit. Descriptive statistics showed that before the intervention, the students in the experimental group generated 0.204 lexical errors per AS-unit and their counterparts in the control group 0.146 lexical errors per AS-unit. After the intervention, the number of lexical errors per AS-unit of the experimental group decreased to 0.156 and that of the control group dropped to 0.140. However, the results of MANCOVA revealed that there was neither significant difference of lexical accuracy between their pre-intervention performance ( $F = 1.022$ ;  $p = 0.317 > 0.05$ ) nor their post-intervention performance ( $F = 0.001$ ;  $p = 0.980 > 0.05$ ).

The results of mixed-design repeated measures ANCOVA indicated that the between-subjects effect was not statistically significant ( $F = 0.339$ ;  $p = 0.563 > 0.05$ ), although graphically the experimental group appeared to have a more salient drop in lexical errors per AS-unit (Figure 6). Simple-effect tests also revealed no significant change over time in either the

experimental group ( $t = 1.333$ ;  $p = 0.193 > 0.05$ ) or the control group ( $t = 0.117$ ;  $p = 0.908 > 0.05$ ).

## Speed Fluency

Speed fluency was estimated by unpruned speech rate (i.e., syllables per minute including all the utterances such as false starts, self-corrections, and repetitions), which was computed by dividing the number of all the syllables produced by the time taken to produce them. Descriptive statistics showed that before intervention, the unpruned speech rate of the students in the experimental group was 147.44 syllables per minute, while that of the students in the control group was 137.25. The MANCOVA results showed that there was no significant difference between the two groups ( $F = 2.555$ ;  $p = 0.116 > 0.05$ ). However, after the intervention, the unpruned speech rate of the students in the experimental group increased to 157.81 and that of the control group students increased to 139.56, indicating a seemingly limited improvement for the control group. The results of MANCOVA revealed that the experimental group outperformed their control group counterparts significantly ( $F = 6.322^*$ ;  $p = 0.015 < 0.05$ ) with a medium to large effect size (partial  $\eta^2 = 0.107$ ) (Cohen, 1988; Miles and Shevlin, 2001).

Likewise, the between- and within-subjects effects were also examined through mixed design repeated measures ANCOVA. A significant between-subjects effect was witnessed ( $F = 5.786^*$ ;  $p = 0.02 < 0.05$ ). With regard to the within-subjects effect, respective simple-effect tests were performed, and the results showed that a significant improvement in speed fluency was seen in the experimental group ( $t = 2.180^*$ ;  $p = 0.038 < 0.05$ ), while the control group did not demonstrate a statistically significant improvement in their speed fluency ( $t = -0.392$ ;  $p = 0.698 > 0.05$ ) (Figure 7).

## DISCUSSION

The results showed that the experimental group students outsourced their control group counterparts on lexical complexity

TABLE 2 | Descriptive statistics.

Metric	Group	Pre-intervention mean	Post-intervention mean	<i>n</i>
G-index	EG	5.60	6.54	29
	CG	5.52	6.10	27
vocd-D	EG	45.36	51.85	29
	CG	38.27	40.14	27
MTLD	EG	32.35	38.07	29
	CG	25.98	28.35	27

EG, experimental group; CG, control group.

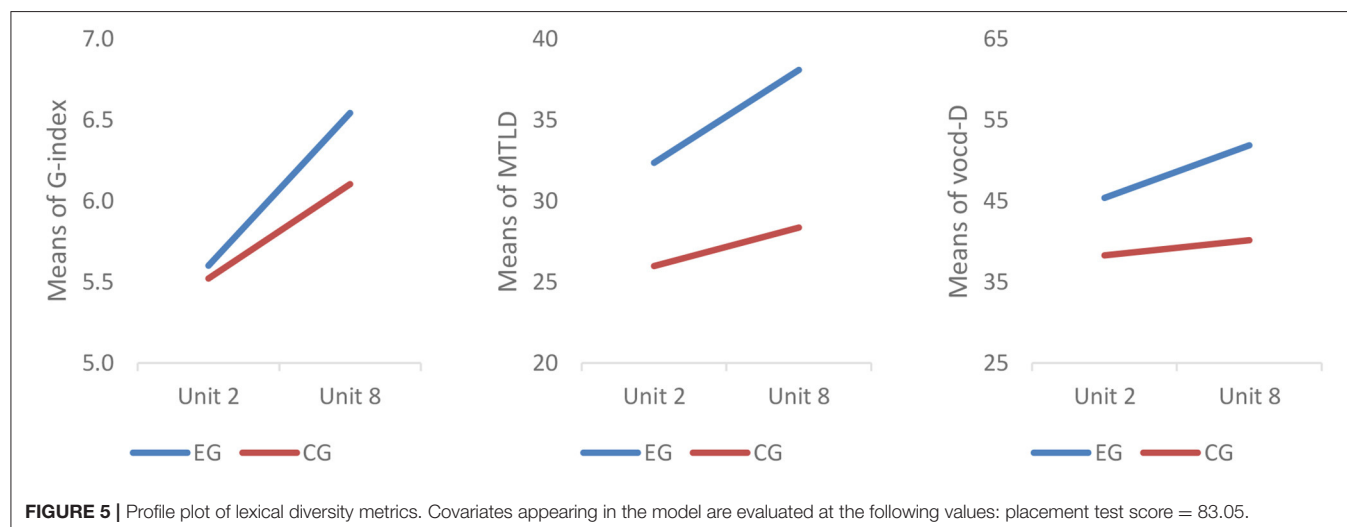
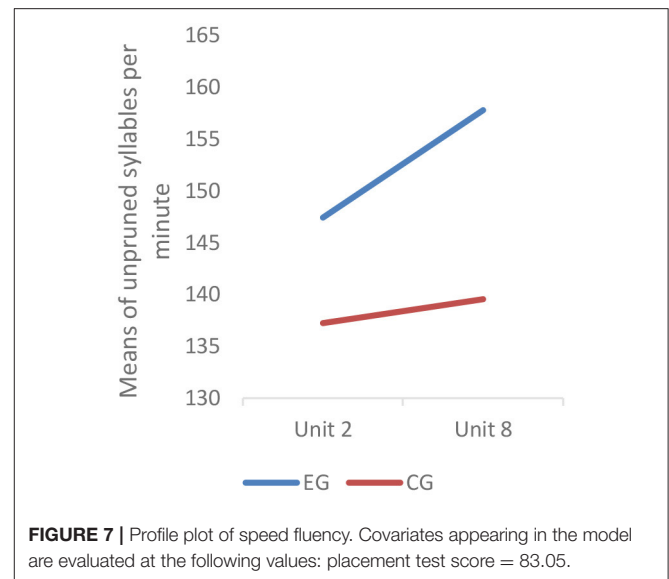
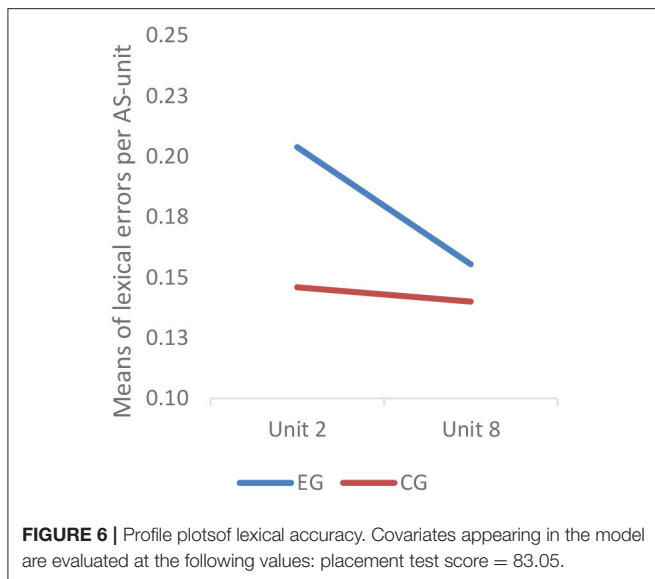


FIGURE 5 | Profile plot of lexical diversity metrics. Covariates appearing in the model are evaluated at the following values: placement test score = 83.05.



(i.e., G-index, vocd-D and MTLT) and speed fluency. But on lexical accuracy, there was no significant post-intervention difference between the two groups. In terms of the within-subjects effect, the experimental group had significant improvement on all three metrics of lexical complexity and speed fluency. In contrast, the control group only had significant improvement on G-index. On lexical accuracy, no significant within-subjects effect was observed in either group.

Generally, the results of this study confirmed the positive effects of integrating the ASR-based application on EFL students' learning, corroborating previous findings (e.g., Evers and Chen, 2020; Dai and Wu, 2021; Jiang et al., 2021). The findings also supported earlier empirical claims that iCALL technologies could provide opportunities for effective vocabulary learning (Chen and Hsu, 2019; Li and Hafner, 2021). It is well-acknowledged that learner preparedness in a flipped classroom plays a pivotal role in students' engagement and task performance in class (Sun and Xie, 2020). Theoretically, owing to the course teachers' pedagogical design, the pre-class self-study in a flipped setting is supposed to be well-organized (Lee and Choi, 2019). However, in practice, students' self-learning is usually affected by factors such as inadequate self-regulated learning ability, resulting in less productive and inefficient preparation for class (Jiang et al., 2020). In the present study, although both groups were learning in a flipped fashion, the significant between-subjects differences indicated that the integration of the ASR-based practice was more goal-oriented and conducive to preparing the students for higher-order interactive tasks in class. Particularly, the ASR-based oral practice featured high interactivity in tandem with synchronic feedback, thus providing the EFL students with ample opportunities to correct themselves. Apart from serving as immediate feedback on the students' utterances, the transcribed texts also visualized students' gradual improvement each time they practiced with self-correction, giving the students an instant sense of accomplishment. It corresponded with the *gradual* feature of effective feedback proposed by Aljaafreh and

Lantolf (1994) that would best promote learners' ZPD from a sociocultural perspective.

Specifically, findings of RQ 1 revealed that benefited from the use of the ASR-based application, the experimental group outscored their counterparts in the control group on G-index, vocd-D, and MTLT in the post-intervention performance, and there was significant within-subjects growth observed on all the three metrics in the experimental group. Measures such as vocd-D and MTLT are deemed critical indicators of L2/FL lexical proficiency, as learners with a richer and more diverse vocabulary are considered as more lexically proficient (Crossley et al., 2009). The experimental group's significant improvement in lexical complexity may be attributed to the ample opportunities for practice made available with the visualized feedback provided by the ASR-based application (Jiang et al., 2021). Before each class, the students in the experimental group performed ASR-enhanced oral tasks which allowed them to correct themselves using the transcribed texts repeatedly. With the aid of the social artifact, they might become more aware of their word choices as they could clearly see every word they uttered while practicing, leading to their deliberate avoidance of a repetitive word used in their previous utterances. In other words, they might vary their use of vocabulary when practicing with the ASR-based application to express themselves. This may well reflect the developmental process from other-regulation to self-regulation as indicated by the sociocultural theory (Wertsch, 1979). In other words, the scaffold afforded by the ASR-based application facilitated the regulation of learners' vocabulary use, which could gradually lead to the internalization of the regulation so that learners became able to self-regulate their vocabulary choice even when such scaffold was released. In fact, this echoes the premise of self-regulated education that the provision of adequate learning scaffolds is always salient in the course of self-regulated learning (Jong, 2019b; Dong et al., 2020). Moreover, the ASR-based practice created an avenue for students to employ the newly learned words in pre-class self-study and hence consolidated

their retention for later use. In group-based discussion, their group members also used some of these words, which further reinforced their grip on these lexical gains. As was argued by El Majidi et al. (2021), such a cyclic lexical process might enable the students to incrementally build a diverse and rich lexicon. Additionally, improvement in vocabulary also seemed to allow more working memory to retrieve more sophisticated vocabulary, therefore producing more lexically diverse and complex speech.

The findings in response to RQ 2 revealed that the two groups had no significant difference in lexical accuracy and neither group had significant gains of lexical accuracy over time, although descriptive statistics showed that the students in the experimental group made fewer lexical errors in the post-intervention UT. This may be attributable to the dictation nature of the ASR technology used in this study. iFlyRec is a speech-to-text dictation ASR application developed for native speakers. Although it was argued that dictation ASR could be more effective in enhancing students' foreign language oracy when combined with scaffolded activities (Evers and Chen, 2020), since it was not designed for pedagogical purposes, dictation ASR-based applications do not provide as sufficient intended feedback on learners' speech as interactive ASR applications such as Google Assistant (Tai and Chen, 2020). This is one of the demerits of dictation ASR technology. Additionally, the role of body language or human emotion in their speech may not be considered in dictation ASR-based oral practice, although they are indispensable elements in human communication. To bridge the limitations of dictation ASR, future studies may combine the use of iFlyRec together with Google Assistant to see whether the two kinds of ASR technologies could jointly improve EFL learners' oracy and further enhance iCALL-based EFL pedagogy.

Contrastingly, the findings of RQ 3 uncovered that significant between- and within-subjects effects in the experimental students' speed fluency were also witnessed in this study. Since the students in the experimental group were encouraged to repeatedly perform the ASR-based tasks (i.e., a condition of task repetition), their pre-class self-study might result in a solid practice effect, which could further lead to a higher degree of automaticity in their utterances when performing the UT. Following Tavakoli et al. (2016), such automaticity is manifested in flow, continuity and smoothness of speech. The automaticity in learners' oral production also coincided with Vygotsky's concept of internalization, which could be interpreted as transforming cognitive functions that are once performed through sociocultural mediation by artifacts and self into cognitive abilities that can be performed independently (Lantolf and Thorne, 2006). Some empirical studies have evidenced that task repetition served as a factor contributing to oral L2/FL fluency (e.g., Ahmadian and Tavakoli, 2011; Goh, 2017). Specifically, the task repetition in this study is regarded as a condition of content repetition, which according to previous studies, tends to be more advantageous for linguistic fluency at the cost of grammatical accuracy (Patanasorn, 2010). On the other hand, the immediate feedback in its written form

provided by the ASR-based application might have enhanced the experimental group students' preparedness for developing a greater degree of automatization in their performance (DeKeyser, 2001, 2007; Segalowitz, 2010; Jiang et al., 2021). Therefore, the pedagogical intervention of ASR-based oral tasks that enabled sustained practice could assist the students in successfully managing their discourse flow when performing the higher-order in-class tasks.

Furthermore, the findings of this study lent some support to Skehan's Trade-off Hypothesis. The triarchic CAF framework generally defines language proficiency as the complex interplay of the three constructs, i.e., complexity, accuracy, and fluency (Tavakoli, 2016), which may be distinctively manifested under different conditions of L2/FL use. The three constructs may be differentially developed by different types of learners and under different learning conditions (Housen et al., 2012). In the current study, the incorporation of ASR-based application for oral practice led to significant growth in the students' speed fluency, while no significant improvement was observed in their lexical accuracy. The contrasting results may indicate a conflict of attention to form and attention to meaning on the learner's side. To be specific, the participants in this study might focus more on the expression of ideas when performing the group-based discussion, indicating a possible priority of meaning over form in oral speech. The practice effect produced by the ASR-based practice seemingly resulted in a degree of proceduralization based on oral lexical chunks, which further led L2/FL learners to develop a state of automatization (DeKeyser, 2001, 2007; Segalowitz, 2010). "When appropriate lexical chunks are readily available, fewer searches are needed, therefore accelerating the formulation process resulting in greater fluidity in oral production" (El Majidi et al., 2021, p. 13). Therefore, in essence, ASR-based technology is regarded as a tool for enhancing L2/FL learners' meaning-oriented proficiency more than form-oriented, indicating that learners may increase their oral fluency at the cost of accuracy.

On the other hand, as was claimed by Skehan (2009), a sequent trade-off might occur between the form-related constructs, i.e., complexity and accuracy, probably because the students were incapable of paying attention to both constructs simultaneously. This result was in line with the previous studies witnessing trade-offs between complexity and accuracy (e.g., Sample and Michel, 2014; Rashtchi and Yousefi, 2017; Granena and Yilmaz, 2019). In the current study, the UTs were considered higher-order tasks that required comprehensive English proficiency, and following Kim (2015), those complex tasks may demand more attentional resources to content, thus allowing less attention allotted to language forms. Although the tripartite CAF conceptualization of L2 performance has become standardized and widely accepted in task-based language teaching (TBLT) (Bui, 2021), the complicated interplay among them remains underexplored in the field of iCALL. Accordingly, more empirical investigations are desirable on this issue in order to understand the interrelationships between the three constructs and how technologies may affect their interplay.

## CONCLUSIONS AND IMPLICATIONS

The current study investigated the effects of the ASR-based technology on EFL students' vocabulary learning based on a pre- and post-intervention quasi-experiment. It was found that the integration of ASR-based technology resulted in significant between-subjects effects on lexical complexity (i.e., G-index, vocd-D, and MTLT) and speed fluency. Conversely, the between-subjects effect on lexical accuracy was not significant. In terms of the within-subjects effect, the experimental group had significant growth on all the three metrics of lexical complexity and speed fluency, while the control group only had significant improvement on G-index. No significant within-subjects effect was seen in either group on lexical accuracy. Given the improvement in EFL students' speed fluency and lexical complexity while not in their lexical accuracy, Skehan's Trade-off Hypothesis was supported in this study.

Pedagogically, the integration of the ASR technology into a flipped foreign language classroom alters the general notion of pre-class self-study in a flipped foreign language classroom which is primarily passive absorption of factual knowledge through pre-recorded video clips on the learner's side. With the ASR-based application, the pre-class oral practice allows the flipped EFL pedagogy to include an active component which provides immediate feedback for students' self-study, thus making it no longer a passive reception of knowledge. Therefore, the ASR-enhanced oral practice can shed light on the pedagogical design of a flipped foreign language classroom. When in-class time is repurposed for higher order language skills in a flipped classroom, the ASR-based technology can be a useful tool for speaking practice, especially when students have limited opportunities to receive feedback on their speaking performance from proficient or native speakers (McCrocklin, 2019).

## LIMITATIONS

Despite the measurable effects of the ASR-based technology on the participants' linguistic performance, the results of this study should be treated with caution due to the following limitations. First, the participants in this study were only enrolled in one university in Chinese mainland, which might raise concerns with the representativeness of the sample. Therefore, more empirical studies conducted in similar research contexts are needed to examine the effects of ASR technology on EFL learners' vocabulary learning. Meanwhile, given that there are studies reporting no evidence for Skehan's Trade-off Hypothesis, EFL teachers need to be cautious about the trade-off in pedagogical practice (Lan et al., 2018). Second, due to the complicated conceptualization and measuring system of the CAF constructs, CAF studies always raise concerns with the operationalizations of multi-dimensional CAF constructs. For example, lexical density and lexical sophistication might be added to the study as another means of measuring lexical complexity. While so far, no synthesis work has been conducted to scope the studies pertaining to the use of ASR technology in L2/FL learning, follow-up studies may

need to employ more comprehensive metrics to perceive CAF as a dynamic and interrelated set of constantly changing subsystems (Norris and Ortega, 2009) in the domain of iCALL. Third, since the Chinese EFL learning context is crucial in understanding the students' EFL learning behavior and in-class peer interaction, a mixed method approach could be employed in future research to draw a holistic picture of how the factors with respect to the Chinese context such as the Chinese educational practice and the local Chinese culture may influence students' in-class task-based oral performance. Fourth, due to the outbreak of the COVID-19 Pandemic, a delayed post-test was not conducted to explore the delayed effects of the ASR technology on the students' vocabulary learning performance. Future studies are advised to perform delayed test to see whether the use of the dictation ASR application has a long-term effect on EFL learners' vocabulary learning.

## DATA AVAILABILITY STATEMENT

The raw data supporting the conclusions of this article will be made available by the authors, without undue reservation.

## ETHICS STATEMENT

The studies involving human participants were reviewed and approved by Survey and Behavioral Research Ethics Committee, the Chinese University of Hong Kong. The patients/participants provided their written informed consent to participate in this study.

## AUTHOR CONTRIBUTIONS

Material preparation, data collection, and analysis were performed by MY-CJ, MS-YJ, C-SC, WW-FL, NW, BS, and BH. The first draft of the manuscript was written by MY-CJ and BS. All authors contributed to the study conception and design, commented on previous versions of the manuscript, read, and approved the final manuscript.

## FUNDING

This work was funded by Fuzhou University Scientific Research Project Fund (Award No. XRC202203).

## ACKNOWLEDGMENTS

Special thank go to the Chinese University of Hong Kong Stanley Ho Big Data Decision Analytics Research Centre and the Chinese University of Hong Kong Teaching Development and Learning Enhancement Grant awarded to Professor Helen Meng for supporting this interdisciplinary research.



## REFERENCES

- Ågren, M., Granfeldt, J., and Schlyter S (2012). "The growth of complexity and accuracy in L2 French," in *Dimensions of L2 Performance and Proficiency: Investigating Complexity, Accuracy and Fluency in SLA*, eds A. Housen, F. Kuiken and I. Vedder (Amsterdam: John Benjamins), 95–119. doi: 10.1075/llt.32.05agr
- Abdulrahman, T. R., and Jullian, M. H. (2020). Engaging young learners in learning vocabulary: a study on learners' perception. *Akademika* 9, 139–153. doi: 10.34005/akademika.v9i01.805
- Ahmadian, M. J., and Tavakoli, M. (2011). The effects of simultaneous use of careful online planning and task repetition on accuracy, complexity, and fluency in EFL learners' oral production. *Lang. Teach. Res.* 15, 35–59. doi: 10.1177/1362168810383329
- Albert, A. (2011). "When individual differences come into play: the effect of learner creativity on simple and complex task performance," in *Second Language Task Complexity*, ed P. Robinson (Amsterdam: John Benjamins), 239–265. doi: 10.1075/tblt.2.16ch9
- Aljaafreh, A., and Lantolf, J. P. (1994). Negative feedback as regulation: second language learning in the zone of proximal development. *Mod. Lang. J.* 78, 465–483. doi: 10.1111/j.1540-4781.1994.tb02064.x
- Anderson, L. W., and Krathwohl, D. R. (2001). *A Taxonomy for Learning, Teaching, and Assessing: A Revision of Bloom's Taxonomy of Educational Objectives*. New York, NY: Addison Wesley Longman.
- Barrot, J., and Gabinete, M. K. (2021). Complexity, accuracy, and fluency in the argumentative writing of ESL and EFL learners. *Int. Rev. Appl. Linguist. Lang. Test.* 59, 209–232. doi: 10.1515/iral-2017-0012
- Bashori, M., van Hout, R., Strik, H., and Cucchiari, C. (2020). Web-based language learning and speaking anxiety. *Comput. Assist. Lang. Learn.* 1–32. doi: 10.1080/09588221.2020.1770293
- Bashori, M., van Hout, R., Strik, H., and Cucchiari, C. (2021). Effects of ASR-based websites on EFL learners' vocabulary, speaking anxiety, and language enjoyment. *System* 99, 102496. doi: 10.1016/j.system.2021.102496
- Bui, G. (2021). Influence of learners' prior knowledge, L2 proficiency and pre-task planning on L2 lexical complexity. *Int. Rev. Appl. Linguist. Lang. Test.* 59, 543–567. doi: 10.1515/iral-2018-0244
- Bulté, B., and Housen, A. (2012). "Defining and operationalising L2 complexity," in *Dimensions of L2 Performance and Proficiency: Investigating Complexity, Accuracy and Fluency in SLA*, eds A. Housen, F. Kuiken and I. Vedder (Amsterdam: John Benjamins), 21–46. doi: 10.1075/llt.32.02bul
- Bulté, B., and Roothoof, H. (2020). Investigating the interrelationship between rated L2 proficiency and linguistic complexity in L2 speech. *System* 91, 102246. doi: 10.1016/j.system.2020.102246
- Chambers, F. (1997). What do we mean by oral fluency? *System*. 25, 535–544. doi: 10.1016/S0346-251X(97)00046-8
- Chavez, M. (2014). Variable beliefs about the need for accuracy in the oral production of German: an exploratory study. *Int. J. Appl. Linguist.* 24, 97–127. doi: 10.1111/ijal.12029
- Chen, H. J. H., and Hsu, H. L. (2019). The impact of a serious game on vocabulary and content learning. *Comput. Assist. Lang. Learn.* 33, 811–832. doi: 10.1080/09588221.2019.1593197
- Cohen, J. (1988). *Statistical Power Analysis for the Behavioral Sciences*, 2nd Edn. Erlbaum.
- Crossley, S., Salsbury, T., and McNamara, D. (2009). Measuring L2 lexical growth using hypernymic relationships. *Lang. Learn.* 59, 307–334. doi: 10.1111/j.1467-9922.2009.00508.x
- Dai, Y. J., and Wu, Z. W. (2021). Mobile-assisted pronunciation learning with feedback from peers and/or automatic speech recognition: a mixed-methods study. *Comput. Assist. Lang. Learn.* 1–24. doi: 10.1080/09588221.2021.1952272
- deBoer, F. (2014). Evaluating the comparability of two measures of lexical diversity. *System* 47, 139–145. doi: 10.1016/j.system.2014.10.008
- DeKeyser, R. (2001). "Automaticity and automatization," in *Cognition and Second Language Instruction*, ed P. Robinson (New York, NY: Cambridge University Press), 125–151. doi: 10.1017/CBO9781139524780.007
- DeKeyser, R. (2007). "Situating the concept of practice," in *Practicing in a Second Language: Perspectives From Applied Linguistics and Cognitive Psychology*, ed R. DeKeyser (New York, NY: Cambridge University Press), 1–18. doi: 10.1017/CBO9780511667275.002
- Dong, A. M., Jong, M. S. Y., and King, R. (2020). How does prior knowledge influence learning engagement? The mediating roles of cognitive load and help-seeking. *Front. Psychol.* 11, 591203. doi: 10.3389/fpsyg.2020.591203
- El Majidi, A., de Graaff, R., and Janssen, D. (2021). Debate as a pedagogical tool for developing speaking skills in second language education. *Lang. Teach. Res.* 1–22. doi: 10.1177/13621688211050619
- Ellis, R. (2003). *Task-Based Language Learning and Teaching*. Oxford: Oxford University Press.
- Ellis, R. (2008). *The Study of Second Language Acquisition*, 2nd Edn. Oxford: Oxford University Press.
- Ellis, R. (2009). Corrective feedback and teacher development. *L2J.* 1, 3–18. doi: 10.5070/L2.V111.9054
- Evers, K., and Chen, S. (2020). Effects of an automatic speech recognition system with peer feedback on pronunciation instruction for adults. *Comput. Assist. Lang. Learn.* 1–21. doi: 10.1080/09588221.2020.1839504
- Foster, P., Tonkyn, A., and Wigglesworth, G. (2000). Measuring spoken language: a unit for all reasons. *Appl. Linguist.* 21, 354–375. doi: 10.1093/applin/21.3.354
- Franco, H., Bratt, H., Rossier, R., Rao Gadde, V., Shriberg, E., Abrash, V., et al. (2010). EduSpeak®: a speech recognition and pronunciation scoring toolkit for computer-aided language learning applications. *Lang. Test.* 27, 401–418. doi: 10.1177/0265532210364408
- Goh, C. C. (2017). Research into practice: scaffolding learning processes to improve speaking performance. *Lang. Teach.* 50, 247–260. doi: 10.1017/S0261444816000483
- Granena, G., and Yilmaz, Y. (2019). Phonological short-term memory capacity and L2 oral performance. *J. Second Lang. Stud.* 2, 317–335. doi: 10.1075/jsls.19005.gra
- Guiraud, P. (1960). *Problèmes et Méthodes de la Statistique Linguistique*. Paris: Presses universitaires de France.
- Han, Y. X., Zhao, S., and Ng, L. L. (2021). How technology tools impact writing performance, lexical complexity, and perceived self-regulated learning strategies in EFL academic writing: a comparative study. *Front. Psychol.* 12:752793. doi: 10.3389/fpsyg.2021.752793
- Hayes, A. F., and Krippendorff, K. (2007). Answering the call for a standard reliability measure for coding data. *Commun. Methods Meas.* 1, 77–89. doi: 10.1080/19312450709336664
- Housen, A., and Kuiken, F. (2009). Complexity, accuracy, and fluency in second language acquisition. *Appl. Linguist.* 30, 461–473. doi: 10.1093/applin/amp048
- Housen, A., Kuiken, F., and Vedder, I. (2012). "Complexity, accuracy and fluency: definitions, measurement and research," in *Dimensions of L2 Performance and Proficiency: Investigating Complexity, Accuracy and Fluency in SLA*, eds A. Housen, F. Kuiken and I. Vedder (Amsterdam: John Benjamins), 1–20. doi: 10.1075/llt.32.01hou
- Jiang, M. Y. C., Jong, M. S. Y., Lau, W. W. F., and Chai, C. S. (2021). Using automatic speech recognition technology to enhance EFL learners' oral language complexity in a flipped classroom. *Aust. J. Educ. Technol.* 37, 110–131. doi: 10.14742/ajet.6798
- Jiang, M. Y. C., Jong, M. S. Y., Lau, W. W. F., Chai, C. S., Liu, K. S. X., and Park, M. (2020). A scoping review on flipped classroom approach in language education: challenges, implications and an interaction model. *Comput. Assist. Lang. Learn.* 1–32. doi: 10.1080/09588221.2020.1789171
- Jong, M. S. Y. (2017). Empowering students in the process of social inquiry learning through flipping the classroom. *Educ. Technol. Soc.* 20, 306–322.
- Jong, M. S. Y. (2019a). To flip or not to flip: social science faculty members' concerns about flipping the classroom. *J. Comput. High. Educ.* 31, 391–407. doi: 10.1007/s12528-019-09217-y
- Jong, M. S. Y. (2019b). Sustaining the adoption of gamified outdoor social enquiry learning in high schools through addressing teachers' emerging concerns: a three-year study. *Br. J. Educ. Technol.* 50, 1275–1293. doi: 10.1111/bjet.12767
- Jong, M. S. Y., Chen, G., Tam, V., Hue, M. T., and Chen, M. (2022). Design-based research on teacher facilitation in a pedagogic integration of flipped learning and social enquiry learning. *Sustainability* 14:996. doi: 10.3390/su14020996
- Jong, M. S. Y., Chen, G. W., Tam, V., and Chai, C. S. (2019). Adoption of flipped learning in social humanities education: the FIBER experience in secondary schools. *Interactive Learn. Environ.* 27, 1222–1238. doi: 10.1080/10494820.2018.1561473
- Kim, I. S. (2006). Automatic speech recognition: Reliability and pedagogical implications for teaching pronunciation. *Educ. Technol. Soc.* 9, 322–334.

- Kim, Y. (2015). "The role of tasks as vehicles for language learning in classroom interaction," in *The Handbook of Classroom Discourse and Interaction*, ed N. Markee (West Sussex: John Wiley and Sons), 163–181. doi: 10.1002/9781118531242.ch10
- Kormos, J., and Dénes, M. (2004). Exploring measures and perceptions of fluency in the speech of second language learners. *System* 32, 145–164. doi: 10.1016/j.system.2004.01.001
- Lambert, C., and Kormos, J. (2014). Complexity, accuracy, and fluency in task-based L2 research: toward more developmentally based measures of second language acquisition. *Appl. Linguist.* 35, 607–614. doi: 10.1093/applin/amu047
- Lan, Y. J., Botha, A., Shang, J. J., and Jong, M. S. Y. (2018). Technology-enhanced contextual game-based language learning. *Educ. Technol. Soc.* 21, 86–89.
- Lantolf, J. P., and Thorne, S. L. (2006). *Sociocultural Theory and the Genesis of Second Language Development*. Oxford: Oxford University Press.
- Lee, J., and Choi, H. (2019). Rethinking the flipped learning pre-class: its influence on the success of flipped learning and related factors. *Br. J. Educ. Technol.* 50, 934–945. doi: 10.1111/bjet.12618
- Lennon, P. (1990). Investigating fluency in EFL: a quantitative approach. *Lang. Learn.* 40, 387–417. doi: 10.1111/j.1467-1770.1990.tb00669.x
- Levis, J. (2007). Computer technology in teaching and researching pronunciation. *Annu. Rev. Appl. Linguist.* 27, 184–202. doi: 10.1017/S0267190508070098
- Li, Y., and Hafner, C. A. (2021). Mobile-assisted vocabulary learning: investigating receptive and productive vocabulary knowledge of Chinese EFL learners. *ReCALL*. 34, 66–80. doi: 10.1017/S0958344021000161
- Liao, J. L. (2020). Do L2 lexical and syntactic accuracy develop in parallel? Accuracy development in L2 Chinese writing. *System* 94:102325. doi: 10.1016/j.system.2020.102325
- Lo, Y. Y., and Murphy, V. A. (2010). Vocabulary knowledge and growth in immersion and regular language-learning programmes in Hong Kong. *Lang. Educ.* 24, 215–238. doi: 10.1080/09500780903576125
- Mackey, A., and Goo, J. (2007). "Interaction in SLA: a meta-analysis and research synthesis," in *Conversational Interaction in Second Language Acquisition*, ed A. Mackey (Oxford: Oxford University Press), 407–452.
- Madini, A. A., and Alshaikhi, D. (2017). VR for teaching ESP vocabulary: a myth or a possibility. *Int. J. Engl. Lang. Educ.* 5, 111–126. doi: 10.5296/ijelev.v5i2.11993
- McCarthy, P. M., and Jarvis, S. (2007) vocd: A theoretical and empirical evaluation. *Lang. Test.* 24, 459–488. doi: 10.1177/0265532207080767
- McCarthy, P. M., and Jarvis, S. (2010). MTL-D, vocd-D, and HD-D: a validation study of sophisticated approaches to lexical diversity assessment. *Behav. Res. Methods* 42, 381–392. doi: 10.3758/BRM.42.2.381
- McCrocklin, S. M. (2016). Pronunciation learner autonomy: the potential of automatic speech recognition. *System* 57, 25–42. doi: 10.1016/j.system.2015.12.013
- McCrocklin, S. M. (2019). ASR-based dictation practice for second language pronunciation improvement. *J. Second Lang. Pronunciation* 5, 98–118. doi: 10.1075/jslp.16034.mcc
- Michel, M. (2017). "Complexity, accuracy and fluency in L2 production," in *The Routledge Handbook of Instructed Second Language Acquisition*, eds S. Loewen and M. Sato (New York, NY: Routledge), 50–68.
- Miles, J., and Shevlin, M. (2001). *Applying Regression and Correlation: A Guide for Students and Researchers*. London: Sage Publications.
- Milton, J. (2013). "Measuring the contribution of vocabulary knowledge to proficiency in the four skills," in *L2 Vocabulary Acquisition, Knowledge and Use: New Perspectives on Assessment and Corpus Analysis*, eds C. Bardel, C. Lindqvist, and B. Laufer (EUROSLA monograph 2), 57–78.
- Mroz, A. (2018). Seeing how people hear you: French learners experiencing intelligibility through automatic speech recognition. *Foreign Lang. Ann.* 51, 617–637. doi: 10.1111/flan.12348
- Nation, I. S. P. (2006). How large a vocabulary is needed for reading and listening? *Can. Mod. Lang. Rev.* 63, 59–82. doi: 10.3138/cmlr.63.1.59
- Neri, A., Mich, O., Gerosa, M., and Giuliani, D. (2008). The effectiveness of computer assisted pronunciation training for foreign language learning by children. *Comput. Assist. Lang. Learn.* 21, 393–408. doi: 10.1080/09588220802447651
- Norris, J. M., and Ortega, L. (2009). Towards an organic approach to investigating CAF in instructed SLA: the case of complexity. *Appl. Linguist.* 30, 555–578. doi: 10.1093/applin/amp044
- Pallotti, G. (2009). CAF: defining, refining and differentiating constructs. *Appl. Linguist.* 30, 590–601. doi: 10.1093/applin/amp045
- Patanasorn, C. (2010). *Effects of Procedural Content and Task Repetition on Accuracy and Fluency in an EFL Contexts*. PhD dissertation, Northern Arizona University, Flagstaff.
- Penning de Vries, B., Cucchiari, C., Bodnar, S., Strik, H., and van Hout, R. (2014). Spoken grammar practice and feedback in an ASR-based CALL system. *Comput. Assist. Lang. Learn.* 28, 550–576. doi: 10.1080/09588221.2014.889713
- Penning de Vries, B., Cucchiari, C., Strik, H., and van Hout, R. (2020). Spoken grammar practice in CALL: the effect of corrective feedback and education level in adult L2 learning. *Lang. Teach. Res.* 24, 714–735. doi: 10.1177/1362168818819027
- Polio, C., and Shea, M. C. (2014). An investigation into current measures of linguistic accuracy in second language writing research. *J. Second Lang. Writing.* 26, 10–27. doi: 10.1016/j.jslw.2014.09.003
- Rahman, A. A., and Angraeni, A. (2020). Empowering learners with role-playing game for vocabulary mastery. *Int. J. Learn. Teach. Educ. Res.* 19, 60–73. doi: 10.26803/ijlter.19.1.4
- Rashtchi, M., and Yousefi, L. M. (2017). Reading input flooding versus listening input flooding: can they boost speaking skill? *J. Lang. Cult. Educ.* 5, 39–58. doi: 10.1515/jolace-2017-0003
- Rassaei, E. (2014). Scaffolded feedback, recasts, and L2 development: a sociocultural perspective. *Mod. Lang. J.* 98, 417–431. doi: 10.1111/j.1540-4781.2014.12060.x
- Rassaei, E. (2020). Effects of mobile-mediated dynamic and nondynamic glosses on L2 vocabulary learning: a sociocultural perspective. *Mod. Lang. J.* 104, 284–303. doi: 10.1111/modl.12629
- Rassaei, E. (2021). Implementing mobile-mediated dynamic assessment for teaching request forms to EFL learners. *Comput. Assist. Lang. Learn.* 1–31. doi: 10.1080/09588221.2021.1912105
- Richards, B. J., and Malvern, D. D. (1997). *Quantifying Lexical Diversity in the Study of Language Development*. Reading: University of Reading New Bulmershe Papers.
- Sample, E., and Michel, M. (2014). An exploratory study into trade-off effects of complexity, accuracy, and fluency on young learners' oral task repetition. *TESL Can. J.* 31:23. doi: 10.18806/tesl.v31i0.1185
- Schmitt, N. (2010). *Researching Vocabulary*. Nottingham: Palgrave Macmillan.
- Segalowitz, N. (2010). *The Cognitive Bases of Second Language Fluency*. New York, NY: Routledge. doi: 10.4324/9780203851357
- Skehan, P. (1996). "Second language acquisition and task-based instruction," in *Challenge and Change in Language Teaching*, eds J. Willis, and D. Willis (Oxford: Heinemann), 17–30.
- Skehan, P. (1998). *A Cognitive Approach to Language Learning*. Oxford: Oxford University Press.
- Skehan, P. (2003). Task-based instruction. *Lang. Teach.* 36, 1–14. doi: 10.1017/S0026144480200188X
- Skehan, P. (2002). Modelling second language performance: integrating complexity, accuracy, fluency, and lexis. *Appl. Linguist.* 30, 510–532. doi: 10.1093/applin/amp047
- Skehan, P., and Foster, P. (2001). "Cognition and tasks," in *Cognition and Second Language Instruction* eds P. Robinson (Cambridge: Cambridge University Press), 183–205. doi: 10.1017/CBO9781139524780.009
- Soleimani, H., Mohammaddokht, F., and Fathi, J. (2022). Exploring the effect of assisted repeated reading on incidental vocabulary learning and vocabulary learning self-efficacy in an EFL context. *Front. Psychol.* 13:851812. doi: 10.3389/fpsyg.2022.851812
- Soyoof, A., Reynolds, B. L., Shadiev, R., and Vazquez-Calvo, B. (2022). A mixed-methods study of the incidental acquisition of foreign language vocabulary and healthcare knowledge through serious game play. *Comput. Assist. Lang. Learn.* 1–34. doi: 10.1080/09588221.2021.2021242
- Sun, B., and Révész, A. (2021). The effects of task repetition on child EFL learners' oral performance. *Can. J. Appl. Linguist.* 24, 30–47. doi: 10.37213/cjal.2021.31382
- Sun, Z. R., and Xie, K. (2020). How do students prepare in the pre-class setting of a flipped undergraduate math course? A latent profile analysis of learning behavior and the impact of achievement goals. *Internet High. Educ.* 46:100731. doi: 10.1016/j.iheduc.2020.100731

- Tai, T. Y., and Chen, H. H. J. (2020). The impact of Google Assistant on adolescent EFL learners' willingness to communicate. *Interactive Learn. Environ.* 1–19. doi: 10.1080/10494820.2020.1841801
- Tai, T. Y., Chen, H. H. J., and Todd, G. (2020). The impact of a virtual reality app on adolescent EFL learners' vocabulary learning. *Comput. Assist. Lang. Learn.* 35, 892–917. doi: 10.1080/09588221.2020.1752735
- Tavakoli, P. (2016). Fluency in monologic and dialogic task performance: challenges in defining and measuring L2 fluency. *Int. Rev. Appl. Linguist. Lang. Test.* 54, 133–150. doi: 10.1515/iral-2016-9994
- Tavakoli, P., Campbell, C., and McCormack, J. (2016). Development of speech fluency over a short period of time: effects of pedagogic intervention. *TESOL Q.* 50, 447–471. doi: 10.1002/tesq.244
- Teng, F. (2019). The effects of video caption types and advance organizers on incidental L2 collocation learning. *Comput. Educ.* 142:103655. doi: 10.1016/j.compedu.2019.103655
- Teng, F. (2020). Vocabulary learning through videos: captions, advance-organizer strategy, and their combination. *Comput. Assist. Lang. Learn.* 35, 518–550. doi: 10.1080/09588221.2020.1720253. [Epub ahead of print].
- Teng, F. (2022). Incidental L2 vocabulary learning from viewing captioned videos: effects of learner-related factors. *System* 105:102736. doi: 10.1016/j.system.2022.102736
- Torlakovic, E., and Deugo, D. (2004). Application of a CALL system in the acquisition of adverbs in English. *Comput. Assist. Lang. Learn.* 17, 203–235. doi: 10.1080/0958822042000334244
- Tweedie, F. J., and Baayen, R. H. (1998). How variable may a constant be? Measures of lexical richness in perspective. *Comput. Humanit.* 32, 323–352. doi: 10.1023/A:1001749303137
- Van Waes, L., and Mariëlle, L. (2015). Fluency in writing: a multidimensional perspective on writing fluency applied to L1 and L2. *Comput. Composition.* 38, 79–95. doi: 10.1016/j.compcom.2015.09.012
- Villamil, O. S., and de Guerrero, M. C. M. (2006). "Socio-cultural theory: a framework for understanding the socio-cognitive dimensions of peer feedback," in *Feedback in Second Language Writing: Contexts and Issues*, eds K. Hyland and F. Hyland (New York, NY: Cambridge University Press), 23–42. doi: 10.1017/CBO9781139524742.004
- Vygotsky, L. S. (1978). *Mind in Society: The Development of Higher Psychological Processes*. Cambridge, MA: Harvard University Press.
- Vygotsky, L. S. (1987). "Thinking and speech," in *The Collected Works of L. S. Vygotsky: Vol. 1: Problems of General Psychology*, eds R. W. Rieber, and A. S. Carton (New York, NY: Plenum), 39–285.
- Wang, Y. H., and Young, S. S. C. (2014). A study of the design and implementation of the ASR-based iCASL system with corrective feedback to facilitate English learning. *Educ. Technol. Soc.* 17, 219–233.
- Webb, S., and Nation, P. (2017). *How Vocabulary Is Learned*. Oxford: Oxford University Press.
- Wertsch, J. V. (1979). "The regulation of human action and the given-new organization of private speech," in *The Development of Self-Regulation Through Private Speech*, ed G. Zivin (New York, NY: John Wiley and Sons), 79–98.
- Wolfe-Quintero, K., Inagaki, S., and Kim, H. (1998). *Second Language Development in Writing: Measures of Fluency, Accuracy, and Complexity*. Honolulu, HI: University of Hawaii Press.
- Xiao, W. Q., and Park, M. (2021). Using automatic speech recognition to facilitate English pronunciation assessment and learning in an EFL context: pronunciation error diagnosis and pedagogical implications. *Int. J. Comput. Assist. Lang. Learn. Teach.* 11, 74–91. doi: 10.4018/IJCALLT.2021070105
- Zhai, X. S., Chu, X. Y., Chai, C. S., Jong, M. S. Y., Istenic, A., Spector, M., et al. (2021). A review of artificial intelligence in education from 2010 to 2020. *Complexity* 2021:8812542. doi: 10.1155/2021/8812542

**Conflict of Interest:** The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

**Publisher's Note:** All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Copyright © 2022 Jiang, Jong, Wu, Shen, Chai, Lau and Huang. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.



## OPEN ACCESS

## EDITED BY

Ehsan Rassaei,  
Majan University College, Oman

## REVIEWED BY

Hung Tan Ha,  
University of Economics Ho Chi Minh  
City, Vietnam  
Lamis Omar,  
Dhofar University, Oman

## \*CORRESPONDENCE

Rui Zhang  
yb57705@connect.um.edu.mo

## SPECIALTY SECTION

This article was submitted to  
Language Sciences,  
a section of the journal  
Frontiers in Psychology

RECEIVED 24 April 2022

ACCEPTED 07 July 2022

PUBLISHED 27 July 2022

## CITATION

Zhang R (2022) Developing  
morphological knowledge with online  
corpora in an ESL vocabulary  
classroom. *Front. Psychol.* 13:927636.  
doi: 10.3389/fpsyg.2022.927636

## COPYRIGHT

© 2022 Zhang. This is an open-access  
article distributed under the terms of  
the [Creative Commons Attribution  
License \(CC BY\)](#). The use, distribution  
or reproduction in other forums is  
permitted, provided the original  
author(s) and the copyright owner(s)  
are credited and that the original  
publication in this journal is cited, in  
accordance with accepted academic  
practice. No use, distribution or  
reproduction is permitted which does  
not comply with these terms.

# Developing morphological knowledge with online corpora in an ESL vocabulary classroom

Rui Zhang<sup>1,2\*</sup>

<sup>1</sup>Research Center for Linguistics and Applied Linguistics, Xi'an International Studies University, Xi'an, China, <sup>2</sup>School of English Studies, Xi'an International Studies University, Xi'an, China

Morphology is the study of word forms and the ways in which words are varied and related to other words in a language. It has been regarded as an essential discipline that is indispensable in language acquisition. It helps learners to figure out the word structure and meaning, particularly the meaning changing of morphemes, which is pivotal for defining words. The present study focuses on developing morphological knowledge with online corpora which are the useful tools for teaching and learning the changes happened in English. Given this light, this research reports an ESL vocabulary classroom in which the instructor designs vocabulary classroom activities with COCA, BNC, TIME MAGAZINE corpus to enrich students' English vocabulary knowledge, help them master word usage, and foster their corpus literacy. This study is oriented by the framework of Classroom Action Research (CAR). Students' classroom performances were recorded and their self-reflections of learning experiences were collected for thematic analysis. The results indicate that, firstly, students' morphological knowledge has been developed as well as their vocabulary knowledge; secondly, their vocabulary self-regulated learning motivation has been incentivised that they are able to answer their own queries about words; thirdly, students' corpus literacy has been fostered, which facilitates their ongoing vocabulary learning; last but not least, a community of inquiry has been established in which students collaborate to construct vocabulary knowledge. The study has wider implications for constructing the student-centered vocabulary classroom and implementing corpus-based instruction in other second language vocabulary classrooms.

## KEYWORDS

**morphological knowledge, online corpora, vocabulary classroom, ESL, community of inquiry, CAR**

## Introduction

Morphology plays an important role in language learning, which deals with how the word forms. In the context of English vocabulary learning, when learners first learn words, they acquire free morphemes, such as “seed”, and they gradually learn about adding a bound morpheme, a suffix “-y”, to form the adjective “seedy”. The meaning varies from a plant having “gone to seed” to “shabby”. This morphological process is involved in the learners' lexical acquisition in which learners examine how the suffixes extend the meaning of roots as well as generate different parts of speech.



As Nagy et al. (1993) point out, students' morphological knowledge of prefixes, roots, and suffixes is fundamental for understanding new words. A lot of research suggests that morphological knowledge facilitates the process of comprehending, storing, and retrieving words (Bowers and Kirby, 2010; Kucan, 2012; Templeton, 2012, 2017; Akbulut, 2017; Goodwin et al., 2020). Therefore, it can be assumed that the more insight the learners have into morphological knowledge, the deeper understandings they can get in their vocabulary learning. Corpora are the useful teaching tools which provide authentic contexts of lexis used in diachronic order. Learners can explore morphological meaning changing and the use of words through concordance lines, whereby they can not only learn the word usage but also explore the historical meaning changing of morphemes. In the morphological process, learners learn vocabulary *via* deducing the word structure and inferring the word meaning with morphological knowledge, and this is endorsed by a couple of researchers that morphological knowledge is a multi-dimension (i.e., semantic, syntactic, phonological, and orthographic information) for learners to explore the nature of words presented as the morphological units (Wolf, 2007; Templeton, 2017; Goodwin et al., 2020).

Many researchers focus on corpus-assisted vocabulary teaching and learning mainly from language production and teaching plan (Frankenberg-Garcia, 2014; Ko and Goranson, 2014; Soruç and Tekin, 2017; Khan, 2019; Rana and Amin, 2020; Ma and Mei, 2021), but few of them concentrate on developing morphological knowledge *via* implementing corpora, which is crucial for vocabulary learning. The present study draws attention to applying online corpora to English vocabulary teaching and learning with the purpose of addressing the following research questions:

1. Is students' morphological knowledge developed with online corpora? What morphological knowledge do they acquire?
2. How do online corpora influence students' vocabulary learning?

The results from students' presentations and self-learning reflections evidently show that students' morphological knowledge has been expanded through exploring the historical meaning changing of morphemes with online corpora. In the vocabulary classroom, students enrich their vocabulary with morphemes and morphemic meanings. Their self-regulated learning motivation has been incentivised that they are encouraged to use online corpora to search words and figure out meanings independently, and they are able to answer their own queries about words *via* the corpora they select. In addition, students' corpus literacy has been fostered, which facilitates their ongoing vocabulary learning. Based on the findings, the present study demonstrates that the corpus-based approach has wider implications for constructing the student-centered vocabulary classroom in ESL and other second language vocabulary classrooms.

## Morphology and corpora

Morphology is known as the study of word formation, focusing on morphemes which are the smallest meaningful units of a language (Anderson, 1982, 1992, 2015; Haspelmath and Sims, 2010; Lieber, 2016). It involves inflection and derivation of morphemes that are closely related to the process of learners' lexical acquisition. As some researchers argue, information about the meaning, pronunciation, and part of speech of a word is derived from its morphological knowledge (Nagy et al., 1993; Stoffelsma et al., 2020; Goodwin et al., 2022). In particular, morphological knowledge is a fundamental dimension for discovering what is in a word, through which learners deduce the word forms and access the morphemes, i.e., prefixes, roots, and suffixes, and these are essential for learners to develop the generic knowledge of vocabulary (Schreuder and Baayen, 1995). When inferring the meaning of a word, morphological meaning changing is crucial since it influences the word formation in ways of derivational regularity and historical borrowing (Anderson, 2015). For example, the suffix “-nik” in English lexis is firstly derived from Russian which means “little”. The word sputnik was found frequently used in the 1950s in the TIME Corpus since the Soviet Union launched the first artificial satellite into space, and it was called Sputnik meaning “little moon” in Russian. Its derivational adjectives such as udarnik (highly-productive shock-brigade worker) and voyenteknik (technician) mean the professionals; beatnik refers to the declined generation of rebellion youth and is related to some kind of fanatic in the 1960s; refusenik indicates a Soviet citizen in the late 1960s when social revolution burst in America, and it was in this period, “-nik” has come to be associated with a sense of communist values. The 1970s onwards show the survival of the main coinages (e.g., kibuzznik) and the arrival of some new ones, but the generative power of the suffix is fading. It can be noticed that languages change as time goes by. During the changing process, new morphemic units come into existence and the established ones may fade away along with the tides of fashion (Baayen, 2009). Such historical meaning changing of morphemes is preferable to be tracked and explored through corpora.

Corpora, which have generally been viewed as a large sample of naturally occurring language, written or spoken, stored in the computer and that can be searched using different kinds of software (Biber et al., 1998; Sinclair, 2005; McEnery and Hardie, 2012; Weisser, 2016; Anderson and Corbett, 2017). Corpora provide a huge number of authentic texts of words and their usage in a real context searched in its context of use, but it does not directly present the meaning of a word which has to be deduced from the concordance lines generated (Q'Keeffe et al., 2007).

The abundance of authentic texts is closely related to the language culture and its linguistic patterns, providing real materials for language learners and this is pivotal for

lexical acquisition (Reppen, 2011). In a corpus, a word and its context are displayed in diachronic order, providing social and historical information about each lexical item—when it first appears, what genre it is applied to, and where it is found. This information, “metadata” (Anderson and Corbett, 2017), is helpful for learners to explore morphology from the aspects of derivational regularities and historical changes, and it is in line with what Stefanowitsch (2020) maintains, that is, ‘corpus morphology is mostly concerned with the distribution of affixes, and retrieving all occurrences of an affix plausibly starts with the retrieval of all strings potentially containing this affix’ (p. 309).

The corpus-based approach is salient for doing diachronic research, offering interesting results when it is used to study language changes. On the bases of corpora, it is palpable that through concordance lines, the context of society, culture, history, and community has been richly glossed, and it is also apparent to identify types containing the affix in concern and review how the words are structured. Like the aforementioned “-nik”, it is transparent to figure out its meaning changing and its combinability through the concordance lines, and it is straightforward to explore the language changes in a natural speech community. By going top down the concordance lines, learners can analyse how the senses shift according to how the words and expressions are used in particular grammatical constructions. Therefore, it can be assumed that using corpora in vocabulary teaching can enable students not only to recognize words but also to analyse their introvert structure systematically and historically with the purpose of facilitating vocabulary teaching and learning.

## Corpus-assisted vocabulary teaching and learning

A number of researchers conducted empirical studies on corpus-assisted vocabulary teaching and learning in the second language learning context (Frankenberg-Garcia, 2014; Ko and Goranson, 2014; Soruç and Tekin, 2017; Khan, 2019; Rana and Amin, 2020). Frankenberg-Garcia (2014) tested the usefulness of separate corpus examples for English comprehension and production among Portuguese secondary school students. The findings show that a single corpus example and multiple corpus examples can help language comprehension. The study also obtains detailed separate encoding and decoding examples to provide evidence in support of corpus-assisted language learning, particularly when learners need different types of examples. However, the question remains to be answered, that is, how to provide learners with easy access to multiple corpus examples mainly on the patterns of language that they need to look up for language production. Soruç and Tekin’s (2017) did a qualitative study investigating 26 learners’ perceptions of corpus-assisted vocabulary learning activities based on COCA and BNC corpora in an EFL classroom in

Turkey. The study collected data *via* reflection papers, semi-structured interviews, and a personal evaluation scale. The results indicated that learners find corpus-assisted vocabulary learning activities interesting, innovative, autonomous, and practical. As Johns (1994) argues, data-driven learning brings learners to authentic language use and corpus rightly holds the feature which helps students develop their learning strategies outside of the classroom with technology. Ko and Goranson (2014), and Rana and Amin (2020) did pre-test and post-test to examine the effectiveness of using corpora in English vocabulary teaching. Their findings suggest that students who are taught by the corpus-based approach performed much better. From their comments, they applaud that corpus provides multiple forms of information for learning new words than guessing the meaning of words from a single context. Moreover, students’ productive knowledge increased significantly, and their vocabulary learning motivation has been strengthened.

Other scholars put forward practical and effective corpus-assisted vocabulary teaching plans. Khan (2019) designed a corpus-assisted vocabulary teaching plan to teach four-word lexical bundles to ESL students at a community college in the USA. The plan suggests that the application of BYU-iWeb corpus for the selection and instruction of four-word lexical bundles helps students attain accuracy and proficiency in their writing expressions. The design enables students to work together by using lexical bundles to produce texts in their academic writing. Ma and Mei (2021) reviewed corpus tools for vocabulary teaching and learning, illustrating how corpora benefit vocabulary teaching and learning and how we choose appropriate corpora. The researchers introduce Corpus-Based Language Pedagogy (CBLP), suggesting designing corpus-based lessons with four design principles—“detecting lexical errors”, “observing and analyzing the language”, “summarizing the language use pattern”, and “practicing using the language”. They also pointed out some issues for teachers to consider in their teaching design—providing students guidance when requiring them to work with corpora, balancing the use of corpora and non-corpora resources, and creating opportunities for students to use the words in context.

As for teaching word formation, a lot of researchers focus on using corpus data to measure morphological productivity and derivation (Baayen and Lieber, 1991; Nagy et al., 1993; Plag and Baayen, 2008; Baayen, 2009; Schröder and Mühleisen, 2010; Fernández-Domínguez, 2013). A few researchers draw attention to implementing corpora in teaching morphology, a pivotal knowledge in vocabulary learning. Wu (2014) discussed the necessity of using corpora in teaching morphology in ESP courses. She argues that corpus data can improve learners’ morphological productivity which in turn assists learners in memorizing ESP vocabulary. Petrovitz and Pierson (2018) analyzed the Chinese English Corpus (CEC) encompassing Chinese international students’ entrance-examination essays. The findings suggest that students have

mastered standard deviational morphology, but simply for correcting language forms and it rarely benefits creativity and effective communication.

Up to this point, there has been little research on the empirical studies of using corpora as a tool for students to develop morphological knowledge, particularly the historical meaning changing of morphemes, which is essential for learning words. The present study demonstrates that online corpora are useful tools for exploring the historical meaning changing of morphemes and this enhances vocabulary teaching and learning.

## Research design

This study is situated in the framework of Class Action Research (CAR), a methodology widely used in the pedagogical field to improve classroom teaching and learning (Burns, 2010; Mettetal, 2012; Mills, 2018). It incorporates informal research practices such as presentations, teaching observation, and learning reflections. Rather than focusing on the statistical or theoretical significance of findings, triangulation of data analysis is performed for validity. By conducting the CAR, students' learning achievements and their learning reflections help teacher-researchers discover the effectiveness of the teaching technique, and this will influence the teacher's teaching refinement which in turn affects students' learning. The characteristic of the CAR takes the form of a dynamic cycle of teaching that brings teaching into a more concrete and efficient orientation than would be the case with purely abstract theorizing. The CAR involved in the present study takes the elements mentioned above into the instructor's consideration: (1) Identifying a research focus. (2) Data Collection. (3) Taking Actions. (4). Reflection.

## Identifying a research focus

In the previous English classroom, the teacher found that, firstly, students have issues in identifying the meaning of words ended up with infrequently used suffixes, such as “-ista”, “-nik”, “-vik”, etc.; secondly, students get panicked when some senses of words are not listed in the dictionary. For example, “seedy”, the “sexually unsavoury” element of its meaning has not found its way into the online *Oxford English Dictionary* where the closest senses are “shabby” or “ill as a result of excessive drink”. Moreover, students are not activated in vocabulary learning. In other words, they are reluctant to develop the knowledge of lexis which is important for lexical acquisition. As Sinclair (1991), one of the pioneers of modern corpus linguistics, noted that ‘the language looks rather different when you look at a lot of it at once.’ (p. 100). Based on these preliminary observations, the present research then focuses on using online corpora to facilitate students' morphological

learning and vocabulary enhancement, and to invoke their English vocabulary independent learning motivation.

## Data collection

Data were collected from 31 students, third-year undergraduates, majoring in English. Among them, 12 are male and 19 are female. The class is for helping them effectively acquire English vocabulary. The author is the teacher-researcher teaching the class. The students attend the course voluntarily. They all signed the consent form. They know that they are engaged in the study and they have the right to end their participation in the research at any time without penalty. The Vocabulary Self-collection Strategy (VSS) chart is used to list the word selection process by students. VSS can promote students' vocabulary learning interest through students' participation in a learning community, in which students share knowledge with words and strengthen self-efficacy in vocabulary learning. According to Haggard (1982), by implementing VSS, students are encouraged to find words and determine the meaning by themselves, and in the next-day presentation, students need to tell where the words are found, what the context-derived meanings are, and why they think the class should know the words. Other students and the teacher discuss about information to reach an agreement on the meanings. After the discussion, a word list is accomplished, and students will have a vocabulary journal at the end of class for review. Students' learning reflections are used to record their learning experience about using online corpora in their vocabulary learning activities.

## Taking classroom action

The vocabulary classroom action is taken according to the following procedure adapted from Yanto and Nugraha (2018):

### 1. Scaffolding

The instructor primarily trains students to use online corpora, such as the Corpus of Contemporary American English (COCA), the TIME Magazine Corpus, and British National Corpus (BNC). After the corpus training, the instructor introduces VSS to students and guides them to make VSS charts with the purpose of initiating students to know about the vocabulary strategy and familiarizing them with the steps: (a) select words from the reading materials, (b) reasons for the selected words, (c) students' definitions of the words, (d) corpus' information of the words and usage. The VSS chart is presented in Table 1.

### 2. Group discussion

Students are divided into small groups consisting of 2–3 students each. In the group discussion, students discuss the meanings of

TABLE 1 VSS chart sample.

No	Word	Source of the word	Reasons for selection	Students' definition	Corpus' information	Word usage
1						
2						
3						

TABLE 2 Student group's VSS chart.

No	Word	Source of the word	Reasons for selection	Students' definition	Corpus's information	Word usage
1	Hemophilia	News	Medical terminology, illness	Blood disorder	The royal disease; effects of the bleeding disorder, including frequent hemorrhages and debilitating pain	Noun and no plural forms; collocated with words like transmit, dread, suffer, etc.
2	Pedophilia	News	Same suffix but the negative meaning	Abnormal hobby	the upsurge of cracking down on internet pedophile pornography crimes	Frequently used with the words like crime, church, study, etc.
3	RUSSOPHILIA	News	Political tendency	People who are friendly toward Russia or fond of Russia and Russian things	Someone who is sympathetic to the political system and customs of the former Soviet Union	Collocated with Trump, flavor of, communist, etc.

words, the reasons for choosing these words, and the importance of words for comprehending the reading content. They show the teacher reading content and the words that they find interesting or hard to understand, and then complete the VSS chart with their group members. Each group offers 2–3 words and the whole class has around 35 words to learn in total.

### 3. In-class presentation

Each group selects a presenter to do the class presentation with PPT slides, reporting the information about their selected words. Other students and the teacher together have a discussion and record the selected words and the contextual meanings in the VSS class chart. The whole class then has an integrated word list for vocabulary learning. Table 2 is one of the groups' VSS charts.

The group presenter shared her group attainment in class with the PPT slides about their learned words. They came across the word “*Socialphobia*” in the web news. When looking for the meaning of “-*phobia*”, the students find another suffix “-*philia*” which has a similar form but opposite meaning to “-*phobia*”. They looked up “-*philia*” in the *Online Etymology Dictionary*, noticing that it is from Greek, meaning “friendship, fondness, tendency toward” and in recent use “abnormal attraction to”. The students are curious to know the words end up with “-*philia*”, its historical information, and its contextual usage. Accordingly, they used the online corpora, COCA and TIME Magazine Corpus, to explore the information they are interested in. The findings are mainly discussed in Figure 1.

From the Figure 1 in the TIME Magazine Corpus, “*hemophilia*” is the only word that appeared most frequently and there are 34 relevant contexts with the suffix “-*philia*” in the 1930s. It was first introduced to the world as “The Royal Disease” during the reign of Queen Victoria of England. She was a carrier of the hemophilia gene. However, it was her son, Leopold, who suffered the symptom of the bleeding disorder, including frequent hemorrhages and debilitating pain. Unfortunately, Leopold passed the carrier gene to his children who eventually married into the royal families of Russia, Spain, and Germany, extending the condition throughout the European Royal bloodlines. This is the key information to illustrate why “*hemophilia*” connects with royalty in the concordance lines, and it collocates with “*suffer*”, “*dread*”, “*transmit*”, etc.

The word “*pedophilia*” came into being at the end of the nineteenth century and widely appeared in the 2000s as the Figure 2 shows.

Particularly after the 1990s, with the upsurge of cracking down on internet pedophile pornography crimes, “*pedophilia*” has gradually become a hot topic in public. In 2003, the United States successfully filmed a documentary that is related to the most sensational case of child molestation in American history. This documentary undoubtedly makes everyone better understand this topic. Moreover, the development of the internet in the last 10 years of the twentieth century allows the proliferation of child pornography. In 1977, 1990, and



TABLE 3 Classification of students' learning reflections.

Categories (student number)	Examples	Proportion
Morphological awareness (SN = 30)	Morphemes are important historical knowledge of morphemes is fantastic morphemes influence the meaning and the structure of words suffixes are complex but important for the meanings of words knowing morphemes is helpful for memorizing the meanings of words word knowledge has been expanded through learning morphemes	96.77
Vocabulary learning motivation (SN = 28)	Enjoy learning English vocabulary motivated to learn new words competent in memorizing words and their meanings like using online corpora to search words learning activities are interesting will learn more words independently	90.32
Corpus literacy (SN = 29)	Access to authentic language usage gain more information than in the dictionary apply it to discourse analysis convenient to analyse data a basic resource for language research challenging and rewarding expect to know more about corpora	93.55
Community of inquiry (SN = 27)	Share findings and understandings with group members collaborate to search words with corpora encourage to learn more words discussions help me clarify my thinking together to resolve word problems	87.10

1994, the United States promulgated laws to punish pedophilia. Accordingly, it can be found from the concordance lines that “*pedophilia*” frequently appears with “*crime*”, “*sexuality*”, “*accuse of*”, etc. In 1982 and 1990, the Federal Supreme Court found evidence related to pedophilia when handling some cases. Therefore, “*pedophilia*” became a public topic widely and was frequently used after the 1980s. In addition, it is surprisingly found that “*pedophilia*” also collocates with “*treatment*”, and this indicates “*pedophilia*” is also treated as a disease.

From the 1960s to the 1990s, some words are structured in the form of “country’s name + -philia”, for example, “*Anglophilia*”, “*Russophilia*”, “*Europhilia*”, etc., see the searching results from COCA (Figure 3).

Why these kinds of words suddenly emerge in the 1960s? This is also the question asked by other students. After introducing the background information of that period, it is discovered that the social situation rightly accounts for the phenomenon. In the late 1950s, the main strength of the capitalist camp differentiated into a tripartite confrontation as America, Japan, and Europe, while the Socialist camp began to split. Therefore, the international political paradigm took up a multi-polarization mode. Under this circumstance, most countries all over the world had their own aligns, and the words, *Anglophilia*, *Russophilia*, and *Europhilia*, appear timely. These words are used adjacently with “*flavor*”, “*support*”, “*communist*”, “*oozing*”, etc. from the concordance lines.

After the in-class discussion, all the students have got a vocabulary list and a vocabulary journal at the end of the semester for them to review whenever and wherever they want. Most importantly, in the process of exploring the word information *via* online corpora, students learn the words along with developing the morphological knowledge and this is completely different from the traditional vocabulary learning in way of content, mode, and cognition. Students’ learning experiences and achievements in morphology and vocabulary were recorded in their reflective journals.

## Data analysis

Thematic analysis (Braun and Clarke, 2006) was used to identify motifs occurring in the content of students’ learning reflections to demonstrate students’ learning achievements and to account for the usefulness of assisting students’ vocabulary learning and morphological knowledge. Data were first assigned *in-vivo* codes. These *in-vivo* codes were then grouped under patterns, through which the themes were identified by comparing shared meanings and characteristics. The coding was done by the author from 35 texts and there were 27 distinct codes generated, through which 4 themes were identified. Results of each coding were checked by a research assistant to test the reliability of the coding system and the validity of the coding process, in which the codes were appropriately grouped under the thematic categories.

## Results and discussion

### Students’ learning reflections

Thirty-one students completed their learning reflections which were coded under the themes of: (a) morphological awareness, (b) vocabulary learning motivation, (c) corpus literacy, (d) community of inquiry. The classification of students’ learning reflections is shown in Table 3.

TIME Magazine Corpus				SEARCH	CHART	CONTEXT	ACCOUNT
CLICK FOR MORE CONTEXT				EXPLORE NEW FEATURES			
1	1930/03/28	Q	Prince of the Asturias (Crown Prince of Spain) was afflicted from birth with <b>haemophilia</b> , a dread and supposedly incurable disease. When a haemophile receives even the slightest				
2	1931/02/13	Q	in that family. The living members provided an ideal group for research. # <b>Hemophilia</b> is a blood disease. The blood coagulates very slowly. A nosebleed may become				
3	1931/02/13	Q	ooze blood for weeks and weeks, until death ensues. Females never have true <b>hemophilia</b> , although some suffer from ailments which seem like this disease. Only males do				
4	1931/02/13	Q	Crown Prince of Spain. Another was the late Tsarevitch of Russia. # Because <b>hemophilia</b> is exclusively a disease of boys and men, Dr. Birch reasoned simply that there				
5	1931/11/16	Q	are like their mother. They are "carriers" of the dread blood disease <b>haemophilia</b> . When he had read the reports, Alfonso XIII as Head of the House				
6	1931/11/16	Q	bleed to death. Well ex-King Alfonso knows that ex-Queen Victoria Eugenie of Spain transmitted <b>haemophilia</b> to their sickly son. Spain's ex-heir, Alfonso, Prince of Asturias.				
7	1931/11/16	Q	's favor which might perhaps have saved the dynasty. # Only males can suffer <b>haemophilia</b> . Only females can transmit it. Mysterious and incurable, - this rare disease				
8	1931/12/14	Q	learning for sure that she and her sister Maria Christina were "carriers" of <b>haemophilia</b> , the family scourge. Last week he smashed the romance of his big-boned 23-year-old				
9	1932/02/01	Q	is Alfonso XIII of Spain. Though the world is welcome to the knowledge that <b>haemophilia</b> taints his family's blue blood, nobody must suspect that he is financially strapped				
10	1932/10/24	Q	(\$2). Introducing mirth-provoking Detective Rossiter in poison cut. hatchet murder and <b>necrophilia</b> . MURDER IN MARYLAND-Leslie Ford-Farrar & Rinehart (\$2). The murder of a				
11	1933/06/12	Q	merchant, whom the Prince met at a Lausanne sanitarium where he was treated for <b>hemophilia</b> . In the face of his father's bitter opposition to the match, the				
12	1933/11/03	Q	, deathly pale, spindle-shanked and likely to bleed to death from the dread disease <b>hemophilia</b> , supplied Spanish Republicans with one of their best reasons for ousting the Royal Ho				
13	1934/08/20	Q	, 20, youngest son of Alfonso XIII, onetime King of Spain; of <b>haemophilia</b> following a slight automobile accident caused when his sister Beatriz swerved into a wall to				
14	1934/08/20	Q	student at the University of Louvain, Belgium, Gonzalo had been considered free of <b>haemophilia</b> , "curse of the Habsburgs." # Died. Leo Falk Wormser,				
15	1934/08/27	Q	by truck. Officially the Government of dictatorial King Alexander frowns on Sarajevo's bland <b>assassinophilia</b> , but His Majesty's police know better than to try' to thwart such				
16	1934/06/04	Q	. She was never really popular among Spaniards. She brought the King the dread <b>haemophilia</b> (easy bleeding) of her house, bore him a haemophilia heir, a				
17	1934/06/04	Q	King the dread haemophilia (easy bleeding) of her house, bore him a <b>haemophilia</b> heir, a second son who was a deaf mute, finally two whole boys				
18	1934/06/04	Q	finally two whole boys, two fine girls capable of passing on their mother's <b>haemophilia</b> . # Last week, fortnight after his 48th birthday, the rumor spread insistently				

FIGURE 1  
Concordance lines of h(a)emophilia.

2002/02/11	Q	cost the Archdiocese of Boston more than \$10 million to settle just some of the <b>pedophilia</b> suits he faces. A Boston Globe investigation revealed last week that as many as
2002/03/04	Q	, ignore or cover up the sexual abuse of children by its own priests? <b>Pedophilia</b> is not a failing; it is not some imperfect but victimless expression of sexuality
2002/03/04	Q	incapable of nurturing kids or -- this dark thought is rarely expressed -- prone to <b>pedophilia</b> . Rebekah Zincavage, a director at the Boston Nanny Center, which employs 300
2002/04/01	Q	1002107 The Rev. Michael Shanahan is struggling to find the Spanish word for <b>pedophilia</b> . First, he ad libs, referring in Spanish to the "crimes of
2002/04/01	Q	. "Then, seeing the rows of blank faces, he resorts to saying <b>pedophilia</b> with a Spanish accent -- which turns out to be right. A white Roman
2002/04/01	Q	; and that raises one of the most vexing questions surrounding the dark condition of <b>pedophilia</b> : Can it be fixed? # The first problem with treating pedophilia is understanding
2002/04/01	Q	condition of pedophilia: Can it be fixed? # The first problem with treating <b>pedophilia</b> is understanding just what it is. While all pedophiles seek sexual gratification with children
2002/04/01	Q	can be satisfied by sexual behavior directed at anyone of any age. # Hard-core <b>pedophilia</b> is a different and more intractable problem, and, according to Tim Smith,
2002/04/01	Q	as molestation but as early initiation. # What still puzzles behavior experts is why <b>pedophilia</b> exists at all. Although it's true that many child molesters were molested as
2002/04/01	Q	wants to talk about their problems. The priests I met were demoralized by the <b>pedophilia</b> revelations and shocked that other bishops hadn't learned the lessons of Chicago, which
2002/04/01	Q	1002136 Well before the scandals over Catholic priests accused of <b>pedophilia</b> , Americans were beginning to cast a skeptical eye on mainstream religion. In a

FIGURE 2  
Concordance lines of pedophilia.

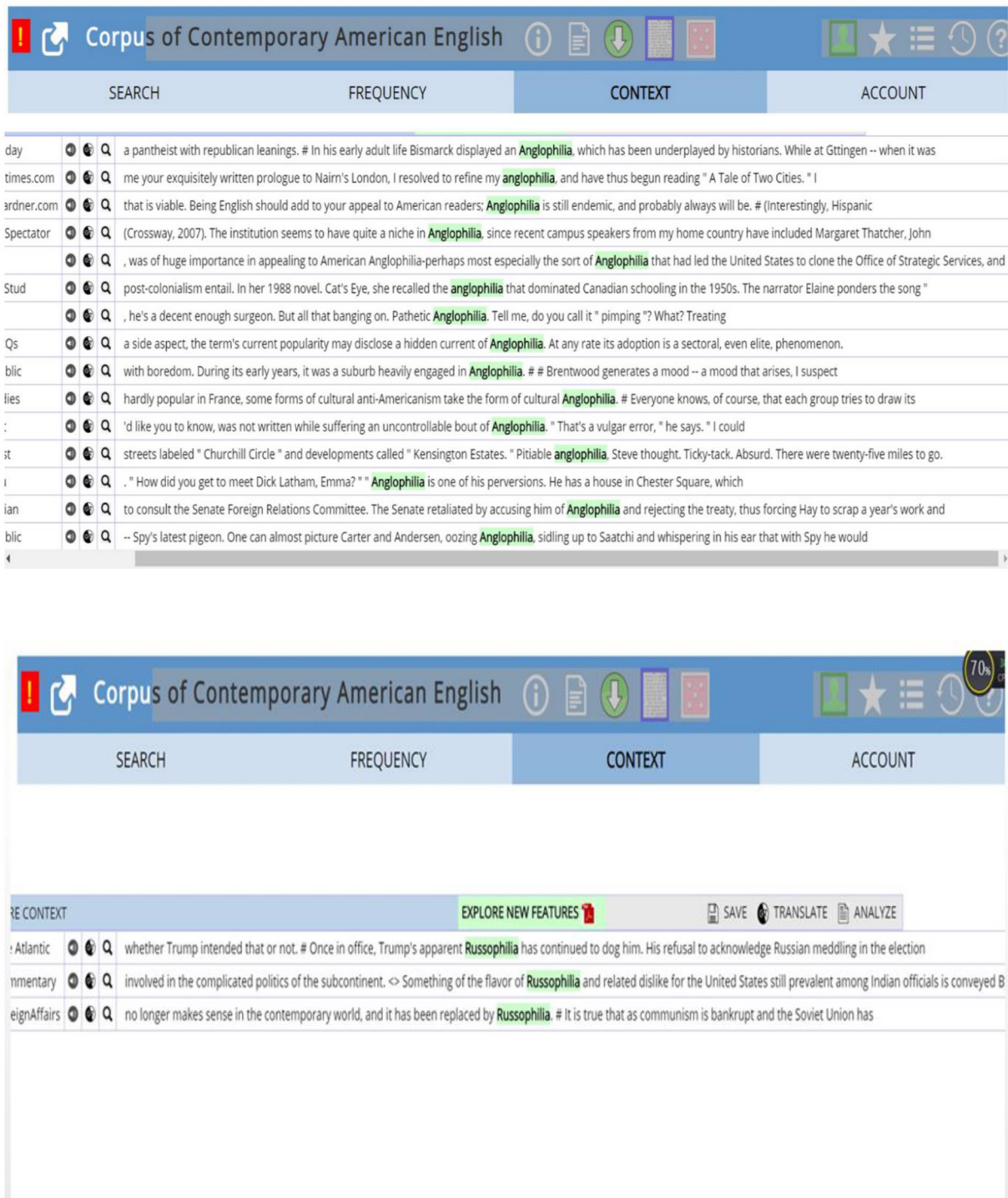


FIGURE 3  
Concordance lines of Anglophilia and Russophobia.

### Morphological awareness

From Table 2, it is noticeable that 96.77 percent of students stated that in the vocabulary learning process,

students became aware of the importance of morphological knowledge, including morphemes and morphological meaning changing when defining the meanings of



words and analyzing the word formation. Excerpts are shown below:

*Student 1. Before, I did not pay attention to morphemes nor pay much attention to the word structure. I just consulted the word meaning from the dictionary and wrote it down. ... Now, I know morphemes are important in vocabulary learning.*

*Student 2. Morphemes, particularly suffixes are complex but now I find them important for figuring out the meanings of words. ... It's easy to use corpora to clarify the meanings of affixes. I will use corpora to look for more words that have the same affixes. It is fun.*

*Students 3. Historical knowledge of morphemes is fantastic. I now understand that it influences the meanings and the structures of words. ... I am now curious about searching for the words that end up with the same suffixes with online corpora and exploring their historical meaning changing.*

*Students 4. There are a lot of morphemes I don't know. I will look for more in my vocabulary learning and search for their meanings with online corpora. ... Knowing morphemes is helpful for memorizing words and meanings. ... My word knowledge has expanded through learning morphemes.*

From the students' reflections, it is evident that students are aware of the importance of morphology in their vocabulary learning, particularly the meaning changing of morphemes. A lot of students stated that historical knowledge of morphemes influences the word structure and meaning. They reflect that it is fun to learn morphemes. As students 2, 3, and 4 mentioned, using online corpora to search words ending up with the same suffixes and exploring the meaning changing of them *de facto* help with defining the words. Moreover, students acknowledge that their word knowledge has been expanded through morphological knowledge. As shown in the in-class presentation, students develop their word knowledge via searching in the online corpora to enlarge their vocabulary stock. In a word, students' morphological awareness has been formed.

## Vocabulary learning motivation

Regarding motivation of vocabulary learning, 90.32 percent of students' learning motivation has been activated and their independent learning has been invoked in way of responding to their own inquiries about vocabulary. Excerpts are displayed as follows:

*Student 5. ... The learning activity is interesting. Making the VSS chart is very helpful for learning words. In the beginning, it's not easy, because finding words and finalizing the word list are time-consuming and skill required. However, it is worth trying. I now enjoy learning English vocabulary.*

*Student 6. In the vocabulary classroom, I feel that I become competent in memorizing words and their meanings. I now like using corpora to search words and their usage. ... My performance in the learning process is good and I did well in the*

*class presentation. I believe that I can do better in the vocabulary learning activities.*

*Student 7. It is fun to learn English words with online corpora. I can find words and understand the meanings by myself. ... Now, when I come up with new words, I can settle them by myself. I can learn more words independently.*

The above reflections reveal that students' English learning motivation has been incentivised. Although some students feel pressure and have difficulties when they use the VSS chart and online corpora for the first time, they eventually become skillful and satisfied with their attainments in their vocabulary learning. It is apparent that students' independent learning has been reinforced that they use corpora to learn words, memorize them, and figure out their meanings and usage by themselves. It can be safely concluded that corpus-based approach pedagogically integrated technology and classroom vocabulary teaching. In addition, students' self-efficacy was fulfilled. They are pleased with their learning performances and believe that they can do better in the vocabulary learning activities. This is in accordance with the research conducted by Zimmerman et al. (1996) that self-regulated learning can enhance self-efficacy. It also demonstrates that self-efficacy is an important element in vocabulary acquisition, influencing independent learning and metacognition which are useful for learners (Mizumoto, 2012).

## Corpus literacy

About 93.55 percent of students show that they accept and fancy online corpora. They are able to search for word information from the authentic language text in the online corpora, and they would like to know and use more corpora to do research. Below are excerpts:

*Student 8. It's a little bit challenging to use online corpora at first, but it is rewarding. I can find more information about words in the online corpora than in the dictionary. ... Online corpora provide various kinds of information, such as collocation, colligation, and data-driven grammar. It enriches my vocabulary knowledge.*

*Student 9. Online corpora open a new world to my English study, allowing me to look at the language usage from quite a novel perspective. The language text that I see from the online corpora is authentic and it is used in an actual context. ... I am intrigued by the corpora and I want to do discourse analysis with them.*

*Student 10. Online corpora show me the material of the language, spoken or written, by the native speakers. It provides a platform where I can gain a better understanding of word usage. ... I now understand MI score can be used to look at the co-occurrences of words and can be compared across corpora. ... I want to know more about corpora to do my research, conversational discourse. It's a basic resource for language research and convenient to analyse data.*

The students' reflections suggest that they are capable to find word usage with online corpora, such as collocation,



colligation, data-driven grammar. They also learned that MI (mutual information) score is used to measure co-occurrences of words and it can be compared across corpora of different sizes. Remarkably, online corpora enable students to access the authentic language used in a real situation and this is the intriguing part that attracts students. Remarkably, some students expect to know more corpora to do research. As students 9 and 10 mentioned, they would like to do discourse analysis and corpora are useful tools for analyzing data. All in all, students' corpus literacy has been well-cultivated, they desire to use corpora not only to enrich their English vocabulary but also to do language research, and this, I argue, is the great effect and achievement of implementing corpora in the vocabulary learning classroom to promote students in their continuous learning.

## Community of inquiry

Results showed that around 87.10 percent of students reckon that in-class presentations and group discussions are beneficial for sharing findings and understandings, collaborating for word search, exploring information, and resolving learning problems. Excerpts are presented as follows:

*Student 11. In the group discussion, we discuss the word list and share our findings and understandings about word information.... I talk to my group members when I don't understand the findings. Discussions help me clarify my thinking when I feel confused.*

*Student 12. I like group work. When I have problems in using online corpora, my group members will help me. We together solve problems.*

*Student 13. In the class presentation, I share my search results with other students and have a discussion with them and the teacher. They ask me some questions and the teacher make some comments about my presentation. ... I have learned a lot.*

The above excerpts notably illustrate that in the process of vocabulary learning, a community of inquiry (Garrison and Vaughan, 2008) has been established, in which students share their findings and understandings of words, settle disputes, gain knowledge about vocabulary, and help each other when coming up with corpus problems. Students, in this vocabulary learning classroom, not only acquire vocabulary knowledge but also cultivate the ability of resolution which profoundly benefits them in their professional study.

As aforementioned, students' learning reflections show that online corpora help with establishing students' morphological awareness. The morphological awareness discussed in this research has shifted the focus from consciousness and ability to the morphemic structure of words aiming to expand and deepen students' word knowledge. It is aligned with Bowers et al. (2010) and (Lyster et al.'s, 2013) argument that morphological knowledge is indispensable to consolidating and refining

students' vocabulary learning. Moreover, by using corpora, students' intrinsic learning motivation is incentivised. They are active to look for words and learn by themselves, and this is better than traditional vocabulary teaching—asking students to look up words in the dictionary and memorize them. As Ryan and Deci (2000) maintain, intrinsic motivation is the inherent tendency to meet the challenges by which learners' learning competence is developed. Therefore, corpus-assisted language learning is an effective approach for teachers to activate students' morphological learning motivation which in turn facilitates their vocabulary acquisition.

Students' learning reflections also indicate that their corpus literacy has been cultivated. They understand what the corpora are, know how to analyse corpora data and how to extract findings from that data. In general, they are motivated to use corpora to learn words and some of them are curious to do language research through corpora. It can be argued that corpora are not only tools for teachers to use in the vocabulary classroom but also a way to arouse students' academic curiosity, through which students are initiated to the linguistic research domain and this manifests the effect of implementing corpora in the language learning classroom. However, students may still need teachers' instructions about corpus utilization, particularly in corpus selection and data sorting. Therefore, teachers should pay more attention to students' corpus-used issues in their learning process to ensure that corpus-based data-driven learning can largely benefit students' language acquisition, and this is also in accordance with Ma et al.'s (2021) point of view that teachers should provide detailed guidance when implementing corpora as resources to address students language learning needs, such as showing the case about selecting corpus, exemplifying data sorting, categorizing results, etc., and in the teaching process, teachers are expected to create sufficient opportunities for students to deal with their target language issues through corpora.

During the collaborative learning process, the community of inquiry has been set up in which students communicate and share their understandings with each other, and develop solutions to problems that can be applied in practice. This collaborative learning environment offers students a learning atmosphere that enables them to negotiate their thinking with shared understandings to construct knowledge. Pedagogically, it is the most rewarding outcome that develops solutions to coursework problems in a community of inquiry (Arbaugh et al., 2008; Garrison, 2017). Overall, based on the results of students' learning reflections, it is apparent that the conduct of corpus-based data-driven learning considerably promotes students' morphological knowledge exploitation, in the process of which students collaborate with each other, make decisions about intended learning words, search for relative information, share and negotiate with peers, present their attainments in class, and this is a cyclical follow supporting their continuous learning.

## Conclusion

This study attests the effectiveness of applying the corpus-based approach in an ESL vocabulary classroom. The praxis is oriented by the framework of CAR. The research findings verify that students' morphological knowledge has been developed and expanded with online corpora. During the learning process, students' morphological awareness has been built. They use morphological knowledge to analyse word structure and infer word meaning, which in turn assist them to memorize words and to practice using them. In this sense, online corpora have enriched students' word inventory and knowledge. Although some students have problems when they first try the online corpora, they soon master them and perform well in the vocabulary activities. They adore using online corpora to search words and plan to use corpora to do discourse analysis, even though doing discourse analysis involves corpus building in another linguistic area—"corpus-based discourse analysis". Notwithstanding, it reveals that students are aware of applying their acquired knowledge in the vocabulary classroom to their English professional study. Furthermore, the community of inquiry has been established in general. Students collaborate to accomplish learning activities, in the process of which they share their understandings, negotiate personal thinking with each other, and construct knowledge. In this way, students' professional knowledge skills are promoted, self-regulated learning motivation is activated, self-efficacy is incentivised, and their professional study expectation is encouraged (Zhang, 2020). Additionally, the study reexamined the practical usefulness of the VSS chart and this could be kept implemented in other second language vocabulary classrooms.

Corpus-assisted vocabulary learning is a new and profound pedagogical approach that modifies EFL/ESL vocabulary classroom teaching. It provides ample authentic language resources for teachers and students to explore morphological knowledge and word formation. However, there are still some limitations when using corpora in the vocabulary classroom. Firstly, some online corpora (e.g., COCA, BYU-BNC, TIME Magazine Corpus, etc.) need an academic license for searching for a number of words. Otherwise, students in classes will be blocked after the total number of queries from the class exceeds 250 searches in a 24-h period, and this might be an issue for universities that have no funds for this expense. Secondly, students might be confused for the first time when exploring the morphological knowledge of vocabulary as there are millions of concordance lines. Teachers can select and tailor the concordance lines for their better understanding.

## Data availability statement

The original contributions presented in the study are included in the article/supplementary material, further inquiries can be directed to the corresponding author/s.

## Ethics statement

The studies involving human participants were reviewed and approved by the Academic Committee of School of English Studies at Xi'an International Studies University. The patients/participants provided their written informed consent to participate in this study.

## Author contributions

The author confirms being the sole contributor of this work and has approved it for publication.

## Funding

This research was supported by the Chinese Educational Ministry Industry—University Cooperative Project under Grant number [202002079007]; Shanghai International Studies University, Institute for Language Materials Development Teaching Materials Research Project under Grant number [2022TX0001]; Xi'an International Studies University Teaching Reform Fund under Grant number [XWK21YB14].

## Conflict of interest

The author declares that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

## Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

## References

- Akbulut, F. D. (2017). Effects of morphological awareness on second language vocabulary knowledge. *J. Lang. Linguist. Stud.* 13, 10–26. Available online at: <https://eric.ed.gov/?id=EJ1140424>
- Anderson, S. R. (1982). Where's morphology? *Linguist. Inq.* 13, 571–612.
- Anderson, S. R. (1992). *A-Morphous Morphology*. Cambridge: Cambridge University Press. doi: 10.1017/CBO9780511586262
- Anderson, S. R. (2015). "Morphological change," in *The Routledge Handbook of Historical Linguistics*, eds C. Bown and B. Evans (New York, NY: Routledge), 264–285.
- Anderson, W., and Corbett, J. (2017). *Exploring English with Online Corpora: An Introduction, 2nd Edn.* Basingstoke: Palgrave Macmillan. doi: 10.1057/978-1-137-43810-2
- Arbaugh, J. B., Cleveland-Innes, M., Diaz, S., Garrison, D. R., Ice, P., Richardson, J. C., et al. (2008). Developing a community of inquiry instrument: testing a measure of the community of inquiry framework using a multi-institutional sample. *Internet Higher Educ.* 11, 133–136. doi: 10.1016/j.iheduc.2008.06.003
- Baayen, H., and Lieber, R. (1991). Productivity and English derivation: a corpus-based study. *Linguistics* 29, 801–843. doi: 10.1515/ling.1991.29.5.801
- Baayen, R. H. (2009). "Corpus linguistics in morphology: Morphological productivity," in *Corpus Linguistics: An International Handbook*, eds A. Ludeling and M. Kyto (Berlin: Mouton de Gruyter), 900–919.
- Biber, D., Conrad, S., and Reppen, R. (1998). *Corpus linguistics: Investigating Language Structure and Use*. Cambridge, UK: Cambridge University Press. doi: 10.1017/CBO9780511804489
- Bowers, P., Kirby, J., and Deacon, H. (2010). The effects of morphological instruction on literacy skills: a systematic review of the literature. *Rev. Educ. Res.* 80, 144–179. doi: 10.3102/0034654309359353
- Bowers, P. N., and Kirby, J. R. (2010). Effects of morphological instruction on vocabulary acquisition. *Read Writ.* 23, 515–537. doi: 10.1007/s11145-009-9172-z
- Braun, V., and Clarke, V. (2006). Using thematic analysis in psychology. *Qual. Res. Psychol.* 3, 77–101. doi: 10.1191/1478088706qp0630a
- Burns, A. (2010). *Doing Action Research in English Language Teaching: A Guide for Practitioners*. New York, NY: Routledge. doi: 10.4324/9780203863466
- Fernández-Domínguez, J. (2013). Morphological productivity measurement: exploring qualitative versus quantitative approaches. *English Stud.* 94, 422–447. doi: 10.1080/0013838X.2013.780823
- Frankenberg-García, A. (2014). The use of corpus examples for language comprehension and production. *Recall* 26, 128–146. doi: 10.1017/S0958344014000093
- Garrison, D. R. (2017). *e-Learning in the 21st Century: A Framework for Research and Practice, 3rd Edn.* New York, NY: Routledge.
- Garrison, D. R., and Vaughan, N. D. (2008). *Blended Learning in Higher Education: Framework, Principles, and Guidelines*. San Francisco, CA: Jossey-Bass. doi: 10.1002/9781118269558
- Goodwin, A., Petscher, Y., and Tock, J. (2020). Morphological supports: investigating differences in how morphological knowledge supports reading comprehension for middle school students with limited reading vocabulary. *Lang. Speech Hear. Serv. Schl.* 51, 589–602. doi: 10.1044/2020\_LSHSS-19-00031
- Goodwin, A., Petscher, Y., Tock, J., McFadden, S., Reynolds, D., Lantos, T., et al. (2022). Monster, P.I.: validation evidence for an assessment of adolescent language that assesses vocabulary knowledge, morphological knowledge, and syntactical awareness. *Assess. Effect. Interv.* 47, 89–100. doi: 10.1177/1534508420966383
- Haggard, M. R. (1982). The vocabulary self-collection strategy: an active approach to word learning. *J. Read.* 26, 203–207.
- Haspelmath, M., and Sims, A. D. (2010). *Understanding Morphology*. (2nd ed.). Routledge.
- Johns, T. (1994). "From printout to handout: grammar and vocabulary teaching in the context of Data-driven learning," in *Perspectives on Pedagogical Grammar (Cambridge Applied Linguistics)*, ed T. Odlin (Cambridge: Cambridge University Press), 293–313.
- Khan, M. A. (2019). New ways of using corpora for teaching vocabulary and writing in the ESL classroom. *ORTESOL J.* 36, 17–24. Available online at: <https://eric.ed.gov/?id=EJ1219789>
- Ko, M. H., and Goranson, J. (2014). Technology-assisted vocabulary learning and student learning outcomes: a case study. *Multimedia Assist. Lang. Learn.* 17, 11–33. doi: 10.15702/mall.2014.17.1.11
- Kucan, L. (2012). What is most important to know about vocabulary? *Read. Teach.* 65, 360–366. doi: 10.1002/TRTR.01054
- Lieber, R. (2016). *Introducing Morphology, 2nd Edn.* Cambridge, UK: Cambridge University Press.
- Lyster, R., Quiroga, J., and Ballinger, S. (2013). The effects of biliteracy instruction on morphological awareness. *J. Immers. Content-Based Lang. Educ.* 1, 169–197. doi: 10.1075/jicb.1.2.02lys
- Ma, Q., and Mei, F. (2021). Review of corpus tools for vocabulary teaching and learning. *J. China Comput. Assist. Lang. Learn.* 1, 177–190. doi: 10.1515/jccall-2021-2008
- Ma, Q., Tang, J., and Lin, S. (2021). The development of corpusbased language pedagogy for TESOL teachers: a two-step training approach facilitated by online collaboration. *Comput. Assist. Lang. Learn.* 4:1–30. doi: 10.1080/09588221.2021.1895225
- McEnery, T., and Hardie, A. (2012). *Corpus Linguistics*. Cambridge, UK: Cambridge University Press. doi: 10.1093/oxfordhb/9780199276349.013.0024
- Mettetal, G. (2012). The what, why and how of classroom action research. *J. Scholarsh. Teach. Learn.* 2, 6–13. Available online at: <https://scholarworks.iu.edu/journals/index.php/josotl/article/view/1589>
- Mills, G. E. (2018). *Action Research: A Guide for the Teacher Researcher, 6th Edn.* Boston, MA: Pearson.
- Mizumoto, A. (2012). Exploring the effects of self-efficacy on vocabulary learning strategies. *Stud. Self Access Learn. J.* 3, 423–437. doi: 10.37237/030407
- Nagy, W., Diakidoy, I., and Anderson, R. C. (1993). The acquisition of morphology: learning the contribution of suffixes to the meanings of derivatives. *J. Read. Behav.* 25, 155–170. doi: 10.1080/10862969309547808
- Petrovitz, W., and Pierson, H. (2018). Morphological creativity in a China English corpus. *Word* 64, 59–68. doi: 10.1080/00437956.2018.1463000
- Plag, I., and Baayen, R. (2008). Suffix ordering and morphological processing. *Language* 85, 109–152. doi: 10.1353/lan.0.0087
- Q'Keeffe, A., McCarthy, M. J., and Carter, R. (2007). *From Corpus to Classroom: Language Use and Language Teaching*. Cambridge: Cambridge University Press. doi: 10.1017/CBO9780511497650
- Rana, A. M., and Amin, Z. (2020). Teaching vocabulary through corpus to the students of graduation. *Int. J. Manage.* 11, 1190–1204. doi: 10.34218/IJM.11.12.2020.111
- Reppen, R. (2011). "Using corpora in the language classroom," in *Materials Development in Language Teaching, 2nd Edn.*, ed B. Tomlinson (Cambridge: Cambridge University Press), 35–50.
- Ryan, R. M., and Deci, E. (2000). Self-determination theory and the facilitation of intrinsic motivation, social development, and well being. *Am. Psychol.* 55: 68–78. doi: 10.1037/0003-066X.55.1.68
- Schreuder, R., and Baayen, R. H. (1995). "Modeling morphological processing," in *Morphological Aspects of Language Processing*, ed L. B. Feldman (Mahwah, NJ: Erlbaum), 131–154.
- Schröder, A., and Mühleisen, S. (2010). New ways of investigating morphological productivity. *Arbeiten aus Anglistik und Amerikanistik* 35, 43–59. Available online at: <https://pub.uni-bielefeld.de/record/1795212>
- Sinclair, J. (1991). *Corpus, Concordance, Collocation*. Oxford: Oxford University Press.
- Sinclair, J. (2005). "Corpus and text - basic principles," in *Developing Linguistic Corpora: A Guide to Good Practice*, ed M. Wynne (Oxford, UK: Oxbow Books), 1–16.
- Soruç, A., and Tekin, B. (2017). Vocabulary learning through data-driven learning in an English as a second language setting. *Educ. Sci. Theory Pract.* 17, 1811–1832. doi: 10.12738/estp.2017.6.0305
- Stefanowitsch, A. (2020). *Corpus Linguistics: A Guide to the Methodology (Textbooks in Language Sciences 7)*. Berlin: Language Science Press.
- Stoffelsma, L., Spooren, W., Mwinlaaru, I., and Antwi, V. (2020). The morphology vocabulary- reading mechanism and its effect on students' academic achievement in an English L2 context. *J. English Acad. Purposes* 2020, 100887. doi: 10.1016/j.jeap.2020.100887

Templeton, S. (2012). "The vocabulary-spelling connection and generative instruction: orthographic development and morphological knowledge at the intermediate grades and beyond," in *Vocabulary Instruction: Research to Practice*, 2nd Edn., eds J. F. Baumann and E. J. Kame'enui (New York, NY: Guilford Press), 116–138.

Templeton, S. (2017). Teaching and learning morphology: a reflection on generative vocabulary instruction. *J. Educ.* 192, 101–107. doi: 10.1177/0022057412192002-312

Weisser, M. (2016). *Practical Corpus Linguistics: An Introduction to Corpus-Based Language Analysis*. Malden, MA: Wiley Blackwell. doi: 10.1002/9781119180180

Wolf, M. (2007). *Proust and the Squid: The Story and Science of the Reading Brain*. New York, NY: Harper/Perennial Press.

Wu, L. (2014). Motivating college students' learning English for specific purposes courses through corpus building. *English Lang. Teach.* 7, 120–126. doi: 10.5539/elt.v7n6p120

Yanto, E. S., and Nugraha, S. I. (2018). Video viewing as a mediation of learning content-based vocabulary: assisting students in understanding disciplinary vocabulary in context. *Indonesian J. Appl. Linguist.* 8, 316–324. doi: 10.17509/ijal.v8i2.13278

Zhang, R. (2020). Exploring blended learning experiences through the community of inquiry framework. *Lang. Learn. Technol.* 24, 38–53. Available online at: <https://doi.org/10.125/44707>

Zimmerman, B. J., Bonner, S., and Kovach, R. (1996). *Developing Self-Regulated Learners: Beyond Achievement to Self-Efficacy*. Washington, DC: American Psychological Association. doi: 10.1037/10213-000





## OPEN ACCESS

## EDITED BY

Ehsan Rassaei,  
Majan University College, Oman

## REVIEWED BY

Hung Tan Ha,  
University of Economics Ho Chi Minh  
City, Vietnam  
Parisa Abdolrezaipour,  
Salman Farsi University of Kazerun, Iran

## \*CORRESPONDENCE

Barry Lee Reynolds  
BarryReynolds@um.edu.mo

## SPECIALTY SECTION

This article was submitted to  
Language Sciences,  
a section of the journal  
Frontiers in Psychology

RECEIVED 13 July 2022

ACCEPTED 01 August 2022

PUBLISHED 02 September 2022

## CITATION

Reynolds BL, Xie X and Pham QHP  
(2022) Incidental vocabulary  
acquisition from listening to English  
teacher education lectures: A case  
study from Macau higher education.  
*Front. Psychol.* 13:993445.  
doi: 10.3389/fpsyg.2022.993445

## COPYRIGHT

© 2022 Reynolds, Xie and Pham. This  
is an open-access article distributed  
under the terms of the [Creative  
Commons Attribution License \(CC BY\)](#).  
The use, distribution or reproduction in  
other forums is permitted, provided  
the original author(s) and the copyright  
owner(s) are credited and that the  
original publication in this journal is  
cited, in accordance with accepted  
academic practice. No use, distribution  
or reproduction is permitted which  
does not comply with these terms.

# Incidental vocabulary acquisition from listening to English teacher education lectures: A case study from Macau higher education

Barry Lee Reynolds<sup>1,2\*</sup>, Xiaowen (Serina) Xie<sup>3</sup> and  
Quy Huynh Phu Pham<sup>4</sup>

<sup>1</sup>Faculty of Education, University of Macau, Taipa, Macao SAR, China, <sup>2</sup>Centre for Cognitive and Brain Sciences, University of Macau, Taipa, Macao SAR, China, <sup>3</sup>Moray House School of Education and Sport, The University of Edinburgh, Edinburgh, United Kingdom, <sup>4</sup>Creative Language Center, Ton Duc Thang University, Ho Chi Minh City, Vietnam

Some proponents of higher education English as a medium of instruction (EMI) have suggested listening to English lectures provides students the opportunity to incidentally acquire unknown words. A case study was designed to examine this assumption. First, the lexical profiles of 27 *Introduction to English Language Teaching* first-year undergraduate course lectures were computed to determine how many words students need to know for comprehension. Then an incoming year-1 undergraduate student with an English vocabulary size of 7,500 word families and mastery of the most frequent 3,000 word families listened to these lectures across 13.5 weeks with the purpose of measuring incidental acquisition gains of three aspects of word knowledge for ten targeted words. Lastly, the student's perceptions about listening to EMI lectures and potentials for this listening inducing incidental acquisition of word knowledge were gathered through a semi-structured interview. The lexical profiling of the entire corpus showed students need knowledge of the most frequent 4,000 English word families plus proper nouns and marginal words for 98% lexical coverage; however, some lectures present students with a more substantial lexical burden than the lectures overall. The student made the most gains in receptive meaning, followed by receptive form, and finally productive meaning. Content analysis of the interview transcript found seven themes representing the student's perception about listening to EMI lectures and their potential for inducing incidental vocabulary acquisition. While the student found listening to the EMI lectures challenging, he perceived the process as useful in preparing for university studies and a career as a secondary English teacher. The student perceived attention, topic, existing vocabulary knowledge, lecturer's native language, and lack of interaction with the lecturer to have moderated incidental learning of vocabulary through listening to English lectures. These results indicate a need to confirm whether incoming students' vocabulary knowledge meet the lexical demands of the EMI lectures given in the Macau context. Furthermore,

pedagogical training on teacher talk strategies and orientation training for incoming students should both be provided to ensure students are receiving high quality instruction.

#### KEYWORDS

incidental vocabulary acquisition, lexical coverage, EMI, listening to lectures, Macau

## Introduction

Extant research on the incidental acquisition of second language vocabulary has mostly focused on the learning of vocabulary that occurs through reading (Reynolds and Wible, 2014). This is because the majority of exposure that second or foreign language learners have to a target language has traditionally been from reading (Nation, 2013a). However, the incidental acquisition of vocabulary from listening has begun to gain traction (e.g., van Zeeland and Schmitt, 2013; Jin and Webb, 2020). This has come along with language learners' increased access to media through the internet and mobile networks (Soyoo et al., 2021). As language learners are exposed to a number of different media, researchers have devoted much attention to understanding the incidental vocabulary learning potentials through learners' consumption of these media. Specifically, researchers have used lexical profiling to show the incidental vocabulary learning potentials from the listening or viewing of television programs, songs, TED talks, YouTube videos, movies, soap operas, and Netflix series, among others (Webb and Rodgers, 2009a,b; Rodgers and Webb, 2011, 2020; Mohsen, 2016; Nguyen and Boers, 2019; Pavia et al., 2019; Fievez et al., 2021; Ha, 2022b).

This line of research has allowed for a deeper understanding of incidental vocabulary acquisition that occurs in informal settings but at an expense of understanding the potentials for incidental vocabulary acquisition occurring in formal settings. Less attention has been given to the incidental exposure to vocabulary that language learners experience when they receive content instruction through the medium of a second language for which the focus of instruction is on learning content and not language. As second language learners in Macau often spend a large number of hours each day receiving content instruction from teachers that use a targeted second or foreign language (Tam and Reynolds, 2022), it is important for researchers to consider the potentials of this exposure for incidental vocabulary learning.

Higher education institutions in Macau are not unlike others around the world that have language policies that promote the use of English as the language of instruction. However, due to the absence of studies in the Macau higher education context, it then remains unknown whether English instructed higher education in the region provides an adequate opportunity for incidental vocabulary acquisition to occur. Thus, a case

study was designed that combined corpus analyses of content course lectures with analyses of a Macau student's incidental vocabulary acquisition and the student's perception of the EMI learning experiences after listening to these course lectures. This case study uncovered the potentials of incidental vocabulary acquisition from listening to teacher talk within the Macau higher education context.

## Incidental vocabulary acquisition from teacher talk

Driven by the proposed benefits brought by the internationalization of education, there is an increasing demand for and assumption that content teachers will use English to deliver their courses (Reynolds and Yu, 2018). The implicit assumption is that receiving content instruction in the English language provides a potential opportunity for students' English vocabularies to develop incidentally through the process of listening to teacher talk (Van Patten, 2003; Horst, 2010). However, only a few studies have tested this assumption empirically within a classroom setting (Smidt and Hegelheimer, 2004; Zhang and Graham, 2020; Brown, 2021).

Examining language related episodes (LRE) occurring inside English medium instruction (EMI) classrooms in two South Korean secondary schools, Hong and Basturkmen (2020) found the majority of the LREs revolved around discipline specific vocabulary. The preoccupation with field specific vocabulary is an indication of its importance to not only understanding content knowledge but also points to the potentials that listening to content lectures have for incidental acquisition of said vocabulary. By contrast, some studies have found listening to teacher talk to be a questionable source for the incidental acquisition of vocabulary (e.g., Horst, 2010; Tang, 2011; Jin and Webb, 2020).

Not all content teachers, especially those from higher education, are willing or able to provide instruction using English. Tang (2011), for example, found the non-native teacher talk of an English teacher in China to be composed of a limited number of words that were also limited in their frequency. These teacher talk phenomena might potentially reduce incidental acquisition outcomes from listening to the lectures. Some teachers may circumvent these issues by increasing the use of first language (L1) translation to encourage the likelihood for incidental acquisition of vocabulary to occur

(Jin and Webb, 2020). In addition, an increase in the frequency of exposure to vocabulary during the lectures can further improve this potential (Dang et al., 2022). By contrast, learners' existing vocabulary knowledge seems to not have as strong of an effect on the likelihood of incidental acquisition. The conflicting results of these studies could be due to their lack of ecological validity as they did not consider incidental learning from listening to lectures for an entire course.

## English as a medium of instruction in Macau

EMI has been a key feature across all levels of Macau's education. Schools in Macau are authorized to choose a medium of instruction. Macau's Education and Youth Development Bureau (DSEDJ) has outlined English curricula for EMI and non-EMI schools (namely, schools with Portuguese or Chinese as the medium of instruction) (Macau Education and Youth Affairs Bureau, 2018). The absence of research on the influences of these EMI and non-EMI curricula on students' English vocabulary development makes it difficult to understand how well students in Macau are prepared by their secondary education for tertiary EMI lectures. These curricula do not specify the expected vocabulary size that students should gain in secondary school and whether this size would be adequate for understanding tertiary level EMI lectures.

Being a multilingual and multicultural region, it is surprising to see a lack of studies looking into how EMI has been implemented in Macau's higher education. One exception is Yu et al. (2020) that explored undergraduate Mainland Chinese students' learning strategies while enrolled in EMI courses in Macau. The university's English for academic purposes courses reported in Yu et al. (2020) did not provide the support needed for the students to excel in the EMI courses. Specifically, the students mentioned the need for more extensive vocabulary instruction. Ma et al. (2022) explored Mainland Chinese postgraduate students' willingness to communicate in English in the Macau EMI classroom. Ma et al. (2022) found students' lack of content knowledge led to a decrease in students' willingness to communicate in the EMI classroom. Rather than fearing that they might lose face as reported in a similar study conducted by Peng (2012), the students were more concerned with their linguistic and content barriers to in-class communication. The findings from these studies indicate that students in Macau are experiencing difficulties in adjusting to the EMI environment.

## Lexical coverage and listening comprehension

Lexical coverage indicates the percentage of words in the input that are known by readers or listeners (e.g., van

Zeeland and Schmitt, 2013; Webb, 2021a). For instance, if students know 95 out of the 100 running words in a given text, the lexical coverage figure would be 95%. In recent years, the relationship between lexical coverage and listening comprehension has been the focus of many studies (e.g., Stæhr, 2009; van Zeeland and Schmitt, 2013; Noreillie et al., 2018; Giordano, 2021). Findings from these studies typically suggest contrasting lexical thresholds for adequate listening comprehension. Stæhr (2009) showed that in the case of advanced Danish learners of English who took the Cambridge English C2 Proficiency exam, 90% coverage (i.e., equivalent to a vocabulary size of the most frequent 2,000 word families) only accounted for 55% comprehension of the listening test. To reach adequate comprehension at an advanced level, 98% coverage was required, meaning that students would need to achieve a vocabulary size of the most frequent 5,000 word families. Noreillie et al. (2018) replicated Stæhr (2009) on a different learner population. They found that at an intermediate level, a vocabulary size of the most frequent 1,250 word families (i.e., 91% coverage) would be sufficient for adequate comprehension. Conflicting evidence regarding lexical thresholds for listening comprehension, as discussed in these two studies, is not surprising because comprehension was measured across learner groups of distinct language proficiencies. Nevertheless, as both studies suggest, a larger vocabulary size leads to more successful listening comprehension.

Analyzing lexical coverage in more informal contexts, van Zeeland and Schmitt (2013) found that 90% coverage was sufficient for adequate comprehension of spoken narratives. However, enormous variation in comprehension at this coverage level was observed, meaning that students' understanding of the texts would be significantly affected by the percentage of the unknown words. At 95% coverage, students were found to achieve a relatively good comprehension. In other words, a vocabulary size of between 2,000 and 3,000 word families would allow them to comfortably understand the spoken texts. Most recently, Giordano (2021) investigated lexical coverage in dialog listening. Participants in the study were at A2-B1 levels based on the Common European Framework of Reference for Languages (CEFR). Students listened to five dialogs at different lexical coverage levels (i.e., 98, 95, 90, 85, and 83%). Similar to the result found in van Zeeland and Schmitt (2013), it was revealed that 90% coverage sufficed for adequate comprehension. Overall, when it comes to lexical coverage necessary for listening comprehension, one would expect a wide range of lexical thresholds from 90% to even 98%, as discussed in previous studies (e.g., Stæhr, 2009; van Zeeland and Schmitt, 2013; Noreillie et al., 2018; Giordano, 2021).

Regarding lexical coverage required for comprehension of academic lectures, to the best of our knowledge, no studies so far have been conducted to examine this relationship. A few studies, however, have examined lexical profiling of academic spoken texts, and the 95 and 98% coverage figures are commonly

cited as the starting point of discussion (e.g., Dang and Webb, 2014; Dang, 2022). For example, Dang and Webb (2014) investigated lexical demands in academic English speech across four disciplines. It was shown that a vocabulary size of the most frequent 4,000 word families plus proper nouns and marginal words allows for 96% coverage, while the knowledge of the most frequent 8,000 word families plus proper nouns and marginal words accounts for 98% coverage. Most recently, Dang (2022) analyzed lexical demands in university lectures across EMI courses, non-EMI courses<sup>1</sup>, and open-access non-EMI courses. It was revealed that non-EMI lectures were the most lexically demanding, while EMI lectures were the least. Specifically, to reach 98% coverage, non-EMI lectures require a lexical demand of the most frequent 9,000 word families, while EMI lectures require a lexical demand of the most frequent 7,000 word families. In the case of open-access non-EMI lectures, a vocabulary size of the most frequent 8,000 word families is needed to reach 98% coverage. Following the current practice on lexical profiling (e.g., Dang and Webb, 2014; Dang, 2022), we also use the 95 and 98% cut-off points to examine lexical profiling of the academic lectures collected for the present study. In other words, we want to examine how many words students would need to know to reach 95 and 98% coverage of these academic lectures.

## Assessing incidental vocabulary acquisition

Acquiring a word through incidental exposure is not an all or nothing phenomenon (Laufer, 1998). Instead, acquisition of different aspects of word knowledge occurs from incidental exposure to language input (Nation and Webb, 2011). This necessitates researchers assessing incidental vocabulary acquisition through tests of varied sensitivity (Nation and Webb, 2011); doing so provides a fuller picture of what aspects of vocabulary knowledge can be acquired from the language input under investigation.

The most influential and thorough taxonomy of word knowledge has been proposed by Nation (2013a). Word knowledge can be dichotomously divided into productive and receptive knowledge. When learners possess receptive knowledge of a word, this refers to their ability to understand, comprehend, or recognize a word while possessing productive knowledge refers to their ability to produce the word (Nation, 2013a). The acquisition of the receptive knowledge of words usually precedes the acquisition of productive knowledge (Zhou, 2010). Nation (2013a) also explains that productive and receptive word knowledge can further be categorized

into three aspects: form, meaning, and use. Form knowledge refers to knowledge of pronunciation, spelling, and word parts. Meaning knowledge refers to the link between form and meaning, a concept and its referents, and the associations that come to mind when a particular word is presented to a learner. Use refers to grammatical functions, collocations, and constraints (e.g., register). A robust incidental vocabulary acquisition research design incorporates measurements of productive or receptive knowledge for each aspect of word knowledge (i.e., form, meaning, and use). Following the current incidental vocabulary acquisition research practices, we assessed both productive and receptive knowledge in addition to three aspects of word knowledge. Assessing these aspects of vocabulary knowledge provides incidental acquisition results on a spectrum of increasing difficulty, namely receptive knowledge of form, receptive knowledge of meaning, and productive knowledge of meaning.

## Problem statement and research questions

While EMI research conducted in the nearby contexts of Mainland China, Hong Kong, Taiwan, Japan, South Korea, and Vietnam have underscored the needs for student support and teacher professional development (Kung, 2013; Hu et al., 2014; Chen and Kraklow, 2015; Aizawa and Rose, 2019; Tri and Moskovsky, 2019; Hong and Basturkmen, 2020; Macaro and Han, 2020; Sung, 2020; Toh, 2020), it is currently unknown whether Macau higher education is also in need of such support. While previous researchers have acknowledged that it is crucial to evaluate the effectiveness of EMI for content learning and English proficiency development (Peng and Xie, 2021), EMI as input for incidental vocabulary acquisition has largely remained under-researched. The studies that do exist lacked ecological validity as they examined incidental vocabulary learning gains from a severely limited number of lectures that were listened to over a very short period of time. More specifically, there is a need to understand the expected English vocabulary size of incoming undergraduates in the Macau region to determine whether they can gain both content and vocabulary knowledge from EMI lectures. Most of the previous EMI studies used either qualitative research methods to understand in-depth views of stakeholders or quasi-experimental designs to compare the learning outcomes of tertiary students (Bowles and Murphy, 2020; Peng and Xie, 2021); however, quantitative data generated from EMI corpus studies are still largely absent in this growing body of EMI research. The use of corpus analyses allow for the potentials of incidental learning of vocabulary to be examined.

To address the research gaps considered above, the present case study investigated the potentials of incidental vocabulary acquisition from listening to teacher talk from

<sup>1</sup> Non-EMI courses include lectures collected in universities in New Zealand, the UK, Ireland, and the U.S (see Dang (2022) for further discussion on non-EMI courses).



EMI course lectures given at a Macau higher education institution. The following research questions guided the study:

1. How many words do students need to know to comprehend the EMI course lectures?
2. How much vocabulary knowledge can the student incidentally acquire from listening to the EMI course lectures?
3. What is the student's perception of the EMI course lectures?

## Materials and methods

The method used received ethical review and approval by the Sub-Panel on Social Science and Humanities Research at the University of Macau (SSHRE22-APP014-FED). A case study was conducted because it allows for close scrutiny of a bounded system (Merriam, 1988). In addition, conducting a case study “lay[s] the groundwork for future studies...[on topics for] which little research has been conducted” (Hood, 2009, p. 70). Case study can illuminate the bounded system and the issue under investigation in a way that can inform whether a future large scale study should be conducted.

## Participants

Two participants were recruited for the case study: a university professor and a secondary school student. Convenience sampling was used to recruit the participants. The university professor (the first author) provided recordings of all 27 lectures from an undergraduate *Introduction to English Language Teaching* course he had previously delivered to first-year students at the University of Macau. The secondary school student was referred to the researchers by his English teacher; he was suitable for the study as he was at the end of his secondary school studies and had just received an admissions offer to the Secondary English Education program in which the undergraduate course is taught.

## Data collection

Several types of data were collected for this case study. These included transcripts of lectures given by the university professor, existing vocabulary knowledge of the secondary school student measured by a vocabulary size and levels test, incidentally acquired vocabulary knowledge of the secondary school student measured by researcher-designed vocabulary tests, and interview data about the secondary school student's

incidental vocabulary acquisition experiences from listening to the lectures.

## Lectures

The lectures were centered on various aspects of language teaching (e.g., native speakerism, intelligibility, instructed second language acquisition, feedback, among others). In 5 out of the 27 lectures, a guest lecturer was invited to interact with the lecturer through a discussion of various topics. The guest lecturer was invited to join several classes because of her knowledge and expertise in the topics under discussion. Regarding her teaching experience and educational background, the guest lecturer has over 10 years of English teaching experience and received her MA and Ph.D. in competitive universities in the United Kingdom. Thus, her English language proficiency would be at C2, the highest level on the Common European Framework of Reference for Languages. At the same time, her speech was perceived as native like by the students enrolled in the course. Accordingly, the inclusion of the guest lecturer would not necessarily affect the analysis of lexical profiling of the current academic lectures. Finally, our corpus data is only made up of academic speech delivered by the lecturers without student interaction. The duration of all the lectures is 15 h and 15 min.

## Vocabulary knowledge

The student's vocabulary knowledge was assessed at the start and end of the case study. At the start, the Vocabulary Size Test (Nation and Beglar, 2007) results indicated the student had a vocabulary size of 7,500 English word families and the Updated Vocabulary Levels Test (Webb et al., 2017) showed mastery of the most frequent 3,000 English word families.

The Vocabulary Size Test is a standardized validated test measuring receptive vocabulary size (Beglar, 2010). For the present study, the student was administered the 140-item version of the Vocabulary Size Test with Chinese options and without an *I don't know* option. While the Vocabulary Size Test is often used in studies that are aimed at measuring breadth of L2 English vocabulary knowledge for the purpose of reading (e.g., Zhang, 2013), it has also been shown to be a significant predictor of L2 proficiency (e.g., Leeming, 2014) and operationalized as L2 proficiency (e.g., Reynolds, 2020). While the student's Vocabulary Size Test results were used as a reference to his English language proficiency, the results of this test have mainly provided information on the student's receptive knowledge of form-meaning connection of written words.

In contrast, the Updated Vocabulary Levels Test measures receptive vocabulary knowledge for five different frequency bands. It can be used to select words for learning or to ensure that learners are working with learning materials or given language input that is at the appropriate lexical difficulty level. Mastery of vocabulary levels was operationalized as answering 29 out of 30 of the items correctly on the 1,000, 2,000, and 3,000

levels and 24 out of 30 of the items correctly on the 4,000 and 5,000 levels of the Updated Vocabulary Levels Test. In other words, the student answered at least 29 of the 30 items on level 1 (30/30), level 2 (30/30), and level 3 (29/30) but was unable to answer 24 out of 30 items on level 4 (23/30) and level 5 (21/30) of the Updated Vocabulary Levels Test. The 29/30 cutoff for the most frequent 3,000 word families of English was suggested by the developers of the Updated Vocabulary Levels test “as these words account for such a large percentage of English [and] they provide the foundation for further lexical and language development” (Webb et al., 2017). It was important to pay close attention to these words because mastery of the most frequent 3,000 word families of English may account for as much as 95% of spoken discourse. Some researchers have claimed mastery of these 3,000 words sufficient for the comprehension of spoken input (van Zeeland and Schmitt, 2013). As the items that appear on the Updated Vocabulary Levels Test and the corpus analyses reported in the present study (see Section “Corpus analyses”) draw on the same lexical sources, using the student’s results on the Updated Vocabulary Levels Test together with the corpus analyses provides an estimate of the words the student was likely to have known while listening to the lectures.

Measuring incidental gains in different aspects of word knowledge required the construction of three assessments at different levels of sensitivity (Nation and Webb, 2011) (see [Supplementary Material](#)). The pre-test and post-test versions of the assessments contained the same items but were presented in a different order to reduce the likelihood of practice effects. Distractors for the receptive form assessment were constructed with the assistance of Wuggy (Keuleers and Brysbaert, 2010), and semantically-related distractors for the receptive meaning assessment were constructed in English before being translated into Chinese, the first language of the student; Chinese distractors were used to ensure that incidental acquisition was being assessed and not comprehension of the items. An *I don’t know* option was included in each meaning recognition assessment item to discourage guessing (Nation and Webb, 2011). The addition of an *I don’t know* option has been shown to reduce the likelihood of guessing (Zhang, 2013). Items on the three word knowledge assessments were marked as either correct or incorrect. The productive meaning assessment was marked independently by the first and second author showing a 100% intermarker agreement. The multiple choice items for the receptive form and receptive meaning assessments were marked using a key.

## Target words

Prior to target word selection, we set a number of criteria. Firstly, given the student’s vocabulary size and knowledge, we selected words that occurred within the 6,000 frequency level and beyond. This reduced the chance that a target word would be known to the student at the start of the study. Secondly, the target words must be content words that appeared at least 5

times in the corpus. Previous research has shown five encounters in auditory input to allow inferring of partial word knowledge (Webb, 2007; Webb and Rodgers, 2009b). Thirdly, selected words must have a relatively similar length in terms of letters and syllables. This will help reduce variation in the difficulty level of the words, thereby improving the reliability of the results. In the end, 10 target words were selected (see [Table 1](#) for a complete profile of the target words). Finally, to ensure that the target words were unknown to the student, we sought professional consultation from his high school teacher, who later confirmed that the student was less likely to know these target words.

## Interview

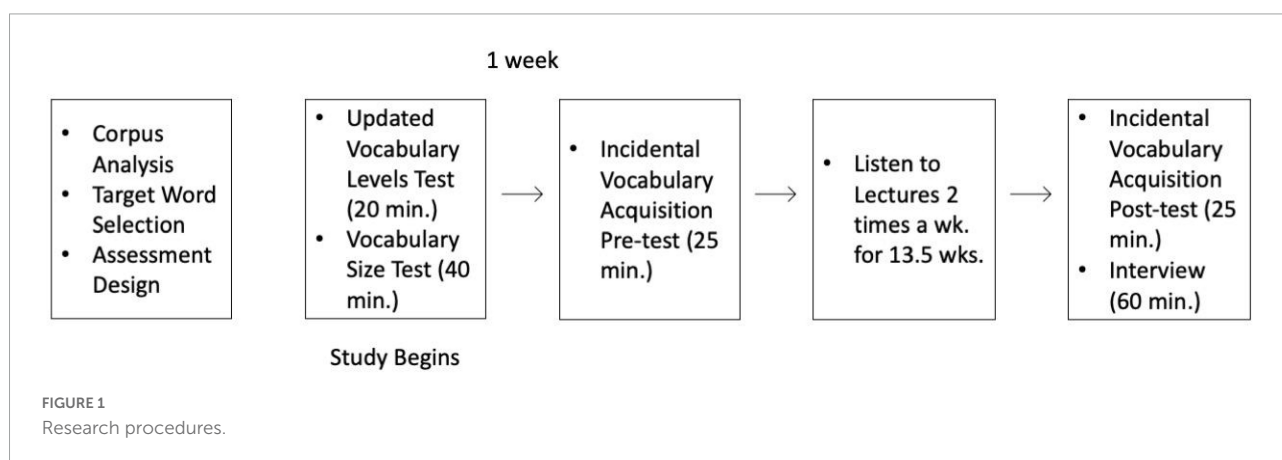
The first author conducted a 60 min semi-structured interview with the student in English (at the student’s request) at the end of the study after he had listened to all of the EMI course lectures (see [Supplementary Material](#)). English was chosen because of the student’s high motivation to engage in English speaking activities. The first author had established a close relationship with the student through bi-weekly contact when providing the student with the recorded lectures. This established comradery made it possible for the student to freely and honestly express his views during the interview.

## Procedures

The procedures are schematized in [Figure 1](#). Prior to the start of the study, the researchers completed the corpus analyses, selected the target words, and designed the vocabulary assessments. Then on the first day of the study, the student was administered the updated vocabulary levels test followed by the vocabulary size test. Then 1 week later, the student was administered the incidental vocabulary acquisition pre-test and then instructed on how to listen to the lectures. The student listened to each lecture a single time (two lectures a week—1 on Friday and one on Sunday) for the next 13.5 weeks. We provided the student access to the lectures one at a time with Google Drive by providing him viewer privileges for a 24 h time period. He was asked to set aside the same time on Fridays and Mondays to listen to each lecture from start to finish without rewinding or fast forwarding. This aimed to simulate a regular course schedule. The student was asked to use the Google Drive web interface to listen to the lectures without rewinding or fast forwarding. He used a Windows 11 compatible desktop computer with a 22" monitor and over-ear headphones to listen to the lectures. At the end of the study an interview was scheduled but on the day of the interview and unbeknownst to the student, the post-test was administered prior to conducting the interview. The post-test was not made known to the student beforehand as this could potentially encourage intentional learning of vocabulary while he listened to the lectures (Nation and Webb, 2011).

TABLE 1 Profile of the target words.

Target words	Part of speech	Letters	Syllables	Frequency of occurrence	Frequency level
Cognate	Noun	7	2	5	13,000
Recast	Verb	6	2	11	10,000
Lexical	Adjective	7	3	35	8,000
Fallacy	Noun	7	3	7	8,000
Syntax	Noun	6	2	8	8,000
Scaffold	Verb	8	2	37	7,000
Normative	Adjective	9	3	6	7,000
Salient	Adjective	7	3	9	6,000
Increment	Noun	9	3	8	6,000
Holistic	Adjective	8	3	5	6,000



## Corpus analyses

To begin the data cleaning process, all the lectures were transcribed using the *Dictate* function in Microsoft Word. Non-verbal actions (e.g., *cough* or *laugh*) were removed since they did not contribute to the lexical demand. Following this was thorough examination of the transcribed lectures to correct misspelling errors made by the transcription software. In the current study, Nation's (2017) British National Corpus (BNC)/Corpus of Contemporary American English (COCA) 25 1,000-word family lists were utilized, with the supplementary four lists of proper nouns, marginal words, compounds, and acronyms. Words that are not found in the wordlists were classified as off-list words. To comply with the spellings used in the wordlists, contractions (e.g., *shouldn't* or *mustn't*) were written in full form, while hyphenated words (e.g., *task-based* or *so-called*) were de-hyphenated. In the case of spelling (e.g., L-E-M-M-A), hyphens were removed, but each letter was written separately (e.g., L E M M A) because it reflected how students perceived the word being spelled in the lectures. Phonetic transcriptions such as *s/or/t/were* were also taken out since students are likely to recognize them due to their high frequency (Dang and Webb, 2014). Furthermore, in our corpus, only five phonetic

transcriptions were found, and thus, the removal of phonetic transcriptions would not significantly affect the analysis. After the data cleaning process, the present corpus consists of 128,498 tokens and 3,952 types. The average speech rate of all the lectures is 140 tokens per minute, which was deemed appropriate for academic lectures, according to Tauroza and Allison (1990).

Prior to data analysis, all lectures were converted into text files. Subsequently, AntWordProfiler (Anthony, 2022) was employed to examine lexical demand of the entire corpus. In the first round of the data analysis, the program identified new proper nouns (e.g., *Moodle* or *WeChat*) and marginal words (e.g., *ahchoo*) that were not listed in the original wordlists. To ensure a more precise calculation of lexical demand of the lectures, these words were added to the original wordlists. In the current study, proper nouns and marginal words were counted in the analysis of lexical coverage on the assumption that they pose a slight learning burden on students (Nation, 2006; Dang, 2020). In other words, it was assumed that students would likely recognize them in lectures. In the case of compounds and acronyms, our data analysis showed that the present corpus contains several uncommon compounds (e.g., *milestone* or *sociocultural*) and an array of professional acronyms (e.g., *CLT* or *CLIL*). Hence, they were excluded from the analysis of lexical

coverage on the grounds that students might be unaware of their meanings. At the same time, the proportions of compounds and acronyms were found negligible, accounting for 0.4 and 0.14%, respectively, and thus, their exclusion was thought not to affect the result. Finally, lexical profiling of each lecture was also reported to provide a more comprehensive picture of lexical demand across different lectures.

The current corpus is made up of academic lectures. Thus, we adopted word family as the counting unit in our study since it is considered as a suitable and useful approach to deal with receptive knowledge (i.e., listening comprehension) (Nation and Webb, 2011; Vilkaitė-Lozdienė and Schmitt, 2019). Following this, we counted the frequency of the target words by word family. In other words, occurrences of all variants of a word would be counted (Reynolds and Wible, 2014). For example, if *colony* occurred 4 times, *colonial* 1 time; *postcolonial* 2 times; *colonizing* 3 times, the cumulative frequency of occurrence of the word *colony* would be 10 times.

## Content analysis

Content analysis (Drisko and Maschi, 2015) was used to analyze the student's responses to the semi-structured interview questions (see [Supplementary Material](#)). Collaboration between co-authors increased the reliability and quality of the analysis (Robson, 2011). The audio recording of the interview was transcribed by a research assistant using the *Dictate* function in Microsoft Word before the research assistant corrected transcription errors. This transcript was then further proofread and corrected by the first and second author. Given the exploratory nature of this qualitative data, inductive coding of the interview transcript was completed in four stages. In stage one, the first and second author became familiar with the data by reading and rereading the transcript. Then, in stage two, the second author loaded the transcript into NVivo Software ver. 1.3.1 to generate initial data-driven codes. In stage three, the first author verified the codes and suggested changes to the second author before the codes were finalized and approved by the first and second author. In the last stage, the codes were combined to form themes that represented the data.

## Results

Our data analysis showed that the majority of tokens (i.e., the occurrence of a word), types (i.e., the unique occurrence of a word regardless of its frequency), and word families (i.e., all variants of a word) were found in the list of the most frequent 1,000 word families. Specifically, 109,952 out of 128,498 tokens, 1,513 out of 3,952 types, and 710 out of 2,190 word families come from the 1,000 frequency level list (see [Table 2](#)). The

finding would come as no surprise given the spoken nature of the current corpus.

In answering RQ1, it was found that a vocabulary size of the most frequent 3,000 word families plus proper nouns and marginal words allows students to understand 97.49% of the entire corpus, while they will need to know the most common 4,000 word families plus proper nouns and marginal words to comprehend 98.02% (see [Table 3](#)). Further analysis, however, showed variations in lexical demand across individual lectures (see [Supplementary Material](#)). Specifically, to reach 98% coverage, nine lectures required a lexical demand of 3,000 word families, five lectures required a lexical demand of 4,000 word families, ten lectures required a lexical demand of 5,000 word families, and one lecture required a lexical demand of 6,000 word families. There were two lectures (lecture 21 and lecture 22 in our corpus) for which a vocabulary size of the most frequent 25,000 word families would only allow for 96–97% coverage. The reason for this is that our analysis excluded compounds which happened to account for a high percentage in these two lectures (3.67% in lecture 21 and 2.15% in lecture 22). Another intriguing finding was that the proportion of proper nouns were particularly large in lecture 9 and lecture 15, accounting for 5.52 and 4.18% in turn, as opposed to their low percentages in the other lectures.

In answering RQ2, the pre-test results showed the student had receptive form knowledge of four targeted words, productive meaning knowledge of zero targeted words, and receptive meaning knowledge of one targeted word. Incidental vocabulary acquisition gains were calculated by subtracting the pretest scores from the posttest scores. The student showed the following *gains* in lexical knowledge: receptive form for two words, productive meaning for one word, and receptive meaning for four words.

Seven themes were generated to answer RQ3. The themes have been placed into three groupings: the student's perception about listening to EMI lectures (see [Table 4](#)); the student's perceptions about listening to EMI lectures for vocabulary gains (see [Table 5](#)); and the student's perceived needs and suggestions for teachers (see [Table 6](#)).

The student claimed two benefits of listening to the EMI lectures. He found the lectures helpful for learning content knowledge (e.g., “th[e] information . . . in these videos are very useful . . . [to] my future . . . lessons later in the university. . . Very useful for those who are going to study English education.) and for gaining vocabulary knowledge (e.g., “I learn[ed] some English vocabulary terms . . . expressions. . .”). By contrast, he mentioned six difficulties in listening to the EMI lectures. As his English lessons in high school differed from the EMI lectures, he found it difficult to stay engaged while listening to the lectures (e.g., “one of the difficult[ies] . . . is to . . . focus on the video for . . . an hour . . .”). He found his language proficiency and content knowledge not sufficient (e.g., “I do not understand the content of [all] those videos . . . I [am] aware . . . that . . . my English is



not ... good enough to ... understand everything in ... those videos...”), thereby making listening to the lectures difficult for him. While he claimed to be a highly motivated student, the potentials of incidental vocabulary acquisition could have been affected by the reduced attention as well as the lack of content and vocabulary knowledge.

The student also held an opinion about the language used by the English native-speaking lecturer. He claimed that the way that his non-native English-speaking teachers gave EMI lectures in high school was different from the way the native-speaking English teacher gave EMI lectures in the recordings. His high school teachers “tend[ed] to ... say everything in an easier way ... [and] those people [in the lectures] sp[o]k[e] in a more difficult way.” He presented a rather contradictory view of the EMI lectures as

containing language that was spoken “in a more difficult way” but could provide him opportunities to “understand more” because the lecturer was willing to explain the new concepts and vocabulary. While he perceived these newly encountered vocabulary as “professional” they were still considered “difficult” as he had never heard them before. He would “guess the meaning from what the person [was] saying,” an example of an incidental vocabulary acquisition strategy. He felt this strategy was not always effective as applying it only resulted in his understanding of around 80% of what was spoken in the lectures. He also shared that his secondary school English teachers did not prepare him to understand university EMI lectures. He claimed “the feeling [of listening to the recorded lectures was] not as familiar as taking lessons in [secondary] school.”

TABLE 2 Tokens, types, and word families at each word level for the entire present corpus.

Word list	Tokens		Types		Word families	
	Raw	%	Raw	%	Raw	%
1,000	109,952	85.57	1,513	38.28	710	32.42
2,000	8,541	6.65	885	22.39	414	18.9
3,000	5,084	3.96	776	19.64	422	19.27
4,000	680	0.53	186	4.71	139	6.35
5,000	757	0.59	99	2.51	78	3.56
6,000	367	0.29	59	1.49	45	2.05
7,000	259	0.2	49	1.24	37	1.69
8,000	96	0.07	32	0.81	29	1.32
9,000	18	0.01	10	0.25	10	0.46
10,000	41	0.03	19	0.48	14	0.64
11,000	47	0.04	15	0.38	14	0.64
12,000	15	0.01	6	0.15	5	0.23
13,000	21	0.02	12	0.3	11	0.5
14,000	21	0.02	7	0.18	5	0.23
15,000	51	0.04	5	0.13	5	0.23
16,000	8	0.01	4	0.1	4	0.18
17,000	37	0.03	5	0.13	5	0.23
18,000	2	0	1	0.03	1	0.05
19,000	0	0	0	0	0	0
20,000	4	0	3	0.08	3	0.14
21,000	0	0	0	0	0	0
22,000	0	0	0	0	0	0
23,000	2	0	1	0.03	1	0.05
24,000	0	0	0	0	0	0
25,000	1	0	1	0.03	1	0.05
Proper nouns	1,589	1.24	145	3.67	129	5.89
Marginal words	90	0.07	10	0.25	8	0.37
Compounds	512	0.4	40	1.01	33	1.51
Acronyms	179	0.14	16	0.4	14	0.64
Off-list words	124	0.1	53	1.34	53	2.42
Total	128,498		3,952		2,190	

The student held perceptions about listening to EMI lectures for vocabulary gains. He perceived the lectures as sometimes enhancing and at other times reducing the opportunities for incidental learning. He shared his opinion on what variables might have influenced the uptake of new vocabulary knowledge. The themes and their corresponding codes are provided in Table 5.

There were two reasons given for why he learned vocabulary from listening to the lectures. First, he mentioned that the words most important to the topic of the lectures were focused on more by the lecturer so this added attention given to them by the lecturer left a lasting impression on him. For example, he shared “another word called Lingua Franca . . . I remember the meaning . . . treating a language as the first language . . . there is Hong Kong English or Singapore English . . . I have gained a lot

more vocabular[y] . . . mostly concept[s].” He was able to recall several words and concepts that he learned about from listening to the lectures. He also shared that he was able to pick up the terms because “the speaker . . . explain[ed] the terms or words very clearly . . . [and] it’s very important to make sure that the students will understand the content.”

The student also shared that he sometimes did not learn new words from listening to the lectures because the words spoken were ones he already knew or he felt were not important to understand the content of the lectures. He would just ignore them. He explained that “there were words that were difficult, but I . . . d[id] not spend a lot of time or effort to . . . get the[ir] meaning because I th[ought they] were not . . . vital to . . . understand . . . the topic that [was] being introduced.”

When further queried on what factors might have affected the incidental learning of vocabulary encountered through listening to the lectures, he said that incidental learning could

TABLE 3 Lexical coverage for the entire present corpus (%) with proper nouns and marginal words.

Word list	Coverage at each 1,000-word level without proper nouns and marginal words	Cumulative coverage with proper nouns and marginal words
1,000	85.57	86.88
2,000	6.65	93.53
3,000	3.96	97.49 <sup>a</sup>
4,000	0.53	98.02 <sup>b</sup>
5,000	0.59	98.61
6,000	0.29	98.9
7,000	0.2	99.1
8,000	0.07	99.17
9,000	0.01	99.18
10,000	0.03	99.21
11,000	0.04	99.25
12,000	0.01	99.26
13,000	0.02	99.28
14,000	0.02	99.3
15,000	0.04	99.34
16,000	0.01	99.35
17,000	0.03	99.38
18,000	0	99.38
19,000	0	99.38
20,000	0	99.38
21,000	0	99.38
22,000	0	99.38
23,000	0	99.38
24,000	0	99.38
25,000	0	99.38
Proper nouns	1.24	
Marginal words	0.07	
Off-list words	0.1	
Tokens	128,498	

<sup>a</sup>Reaching over 95% coverage, <sup>b</sup>reaching over 98% coverage.

TABLE 4 The student’s perceptions about listening to EMI lectures.

Theme	Codes
Benefits of listening to the EMI lectures	Content learning Vocabulary gains
Difficulties when listening to the EMI lectures	Difficulty with engagement while listening Difficulty in understanding due to a lack of language proficiency Difficulty in understanding due to a lack of content knowledge Difficulty in understanding due to native speaker lecturer’s language use Vocabulary and learning strategy used Lack of secondary school English training in understanding EMI lectures

TABLE 5 The student’s perceptions about listening to EMI lectures for vocabulary gains.

Theme	Codes
Possibility for vocabulary gains	Source for vocabulary gains – vocabulary used related to the EMI lecture topics Teacher talk strategy aided vocabulary gains – explanations
Reasons for reduced possibility for vocabulary gains	Easier vocabulary used Not essential vocabulary for understanding content
The perceived influence of vocabulary gains	Teacher’s explanations For future study in an EMI environment For future teaching career

have been affected by the teacher’s explanations of the unknown words, whether he felt the unknown words were useful for his future university studies, and whether the vocabulary would be useful to his future career goal of becoming a secondary school English teacher. He explained that if the lecturer did not “explain very . . . comprehensively, [he would] not have a very good impression about the word.” When he came across words, especially concepts related to English education, his attention to the lectures would be heightened as he felt listening to these particular parts was “getting [him]self more prepared for the things [he] will learn in the university . . . [and he] might need to use those concepts or ideas that are introduced in those videos . . . when [he] becomes . . . [an] English teacher.”

The student also provided suggestions to both his future university lecturers and his previous secondary school English teachers. He provided five pieces of advice for university lecturers and three pieces of advice to secondary school English teachers (see Table 6).

The student said that university EMI lecturers should be aware that the language proficiency “level of students varies.” This means some students may need more scaffolding than others to understand the content. This thought caused the student to consider whether he could have learned more content knowledge from the lectures if the lectures had been delivered in Chinese, his first language. He shared that “it might be more effective . . . if . . . Chinese [was used] to express the meaning, because [he] might have a stronger impression and a better understanding of those vocabulary . . . [This] might make it more . . . efficient and more effective. . .”. The student explained an internal struggle he experienced when thinking

about whether English or Chinese should be used in higher education. He concluded his answer by saying that lectures “in English will [provide] more improvement in . . . English . . . [and lectures] in Chinese . . . might not [result in] good English.”

In order to provide students the opportunity to improve their English through listening to English lectures while also taking care of content knowledge needs, he suggested university lecturers could provide additional reading materials and videos to students to review before listening to lectures. He assumed that when he took courses in the future that the lectures would be accompanied by “passage[s] or book[s]” and that he could “search for the meaning of the words that [he] did not understand to . . . [understand] the content more thoroughly.” The discussion of what might occur in a future EMI classroom led to his sharing that he might have learned more vocabulary if there was interaction among the lecturer and students. For this study, he only listened to the lectures but in a formal classroom he would not only listen to lectures but also have the opportunity to interact with classmates and offer questions to the lecturer for clarification. He felt that listening to lectures for this study:

. . . must be more difficult than having lessons in the classroom because . . . I am just receiving things . . . [and] not able to ask question[s]. I [can’t ask] the speaker to elaborate [on] something or to further explain something . . . that’s one of the shortcoming[s] of [listening to] these lectures.

He felt that in the regular classroom, a lecturer would increase the opportunities for incidental acquisition of vocabulary simply by allowing students to ask questions.

The student also offered three suggestions to secondary school English teachers to help prepare students for the experience of listening to EMI lectures in university. He felt that by sharing with the students what they are going to experience in university, the secondary school English teachers could motivate students to put more effort into their English studies. They could try to provide them with examples of lectures such as the ones he viewed so that students would know what to expect and would not have the same assumptions as he had before this experience. He felt that this could help “the students’ mindsets prepare [for] how it will be like in . . . the university.” The teachers could also try to use such materials as the content in their English classes to gradually get the students familiar with content learning in English.

Discussion

Our findings showed that a vocabulary size of the most frequent 3,000 word families plus proper nouns and marginal words enables students to understand 97.49% of all the present EMI lectures combined. If they know the most frequent 4,000

TABLE 6 The student’s perceived needs and suggestions for teachers.

Theme	Codes
Suggestions for university EMI lecturers	Consider students’ varied language proficiency levels and needs
	Students need to use L1 and translanguaging in EMI courses
	Potential impact of using L1 and translanguaging in EMI courses
	Strategies to aid vocabulary learning – additional readings and videos
	Teacher talk strategies can enhance EMI learning experience – interaction in the classroom
Suggestions for secondary school English teachers	Motivate students to learn English
	Raise students’ awareness about university EMI lectures
	Use readings and videos as teaching materials to prepare students for university EMI lectures

word families plus proper nouns and marginal words, they can understand 98.02%. However, Dang (2022) found that to reach 98% coverage of EMI lectures, students would need to achieve knowledge of the most frequent 7,000 word families. Clearly, there was a huge difference between the lexical demands for our corpus and the corpus presented in Dang (2022). Such a difference can be attributed to many factors. First, the size of our corpus is significantly smaller than the size of the EMI corpus in Dang (2022), 128,498 and 253,906 words, respectively. Clearly, the corpus of a bigger size would contain significantly more new words, resulting in a greater lexical demand, especially at 98% coverage. Secondly, our corpus data is derived from one single discipline, which is *English Language Teaching*; however, the EMI corpus in Dang (2022) consists of various disciplines. This could account for variations in lexical demand across EMI lectures of different disciplines. Another potential factor is the delivery of the lectures. It could be that lexical usage in our corpus was intentionally modified by the lecturers to facilitate students' understanding of the new concepts and ideas. This helps explain why lexical demand at 98% coverage was found to be lower in our study compared to what was found in Dang (2022). Undoubtedly, this claim is tentative, but it provides an important pedagogical implication in that the analysis of lexical demand of academic lectures should be contextualized within a specific classroom. Overall, we want to emphasize that only through the examination of lexical demand of the lectures collected in a specific course can we provide students with a more precise and realistic estimation of how many words they need to know to reasonably understand the lectures. This is because lexical coverage of academic lectures can be affected by various factors, such as the duration of the lectures, the disciplines, or even the form of delivery.

Variations in lexical demand across individual lectures were also observed in our study. Specifically, ten lectures required students to possess a vocabulary size of the most common 5,000 word families to comprehend 98% of the texts, and one lecture required a lexical demand of 6,000 word families to achieve this same level of understanding. Pedagogically, we suggest that in the case of more linguistically complex lectures, teachers should introduce students to novel words in advance (e.g., having them do a vocabulary quiz in class or asking them to read the textbook materials before class). This would maximize students' understanding of the subject content.

Our study excluded compounds from the analysis of the lexical coverage of the current corpus on the grounds that several compounds were found to be advanced lexical items and could create learning difficulties for students. However, analysis of individual lectures showed that two out of the twenty-seven lectures consisted of a high proportion of compounds. Pedagogically, our finding suggests that in the case of infrequent compounds, teachers could introduce them to students beforehand to improve students' understanding of the lesson content. Methodically, our finding raises an important

question in research on lexical coverage, that is, how to best deal with compounds (e.g., whether to include or exclude them from the lexical coverage analysis). Cobb (2019) proposed an interesting approach that compounds should be first broken down (e.g., *shutdown* would be broken into *shut* and *down*), and then their pieces will be put back to the wordlists for analysis. Breaking compounds down, however, would eventually increase the number of words in the corpus, thus affecting the calculation of the speech rate. Clearly, more research on how to effectively deal with compounds in the calculation of lexical coverage is warranted.

In the case of proper nouns, they are usually included in the analysis of lexical coverage based on the assumption that proper nouns pose a slight learning burden on students (e.g., Nation, 2006; Webb and Rodgers, 2009b; Dang, 2020). In our study, the proportion of proper nouns accounted for 1.24% of the entire corpus. However, this percentage is immensely large in certain lectures. Specifically, the proportion of proper nouns is even higher than the coverage of the 3,000-word level list in lectures 9 and 15. Kobleva (2012) found that unfamiliar proper names can significantly affect students' listening comprehension. Pedagogically, we recommend that if a particular lecture contains a large number of unfamiliar proper nouns, teachers should pre-teach them to aid students' understanding of the lectures. Likewise, it might be unreasonable to expect students without mastery of the most frequent 3,000 words of English to comprehend lectures from this course. Methodologically, our study raises an important question as to whether unfamiliar proper nouns should be included in the calculation of lexical coverage given that they can pose difficulties to students when they listen to academic lectures.

The student's vocabulary knowledge should have provided him almost 98% lexical coverage of the lectures. This seems ideal for encouraging incidental vocabulary acquisition; however, after listening to 15 h and 15 min of lectures, the student gained receptive knowledge of form for two words, productive knowledge of meaning for one word, and receptive knowledge of meaning for four words. These results may seem very modest but it must be noted that the student is likely to have acquired knowledge of words that were not targeted in the present study. Certainly, various factors came into play and affected the incidental vocabulary learning from listening to academic EMI lectures. While it is difficult to pinpoint these variables, frequency of exposure and salience are likely to have contributed to the incidental acquisition of word knowledge.

Frequency of exposure has been shown to be a robust predictor of incidental vocabulary acquisition (e.g., Rott, 1999). The frequency effect seems to have resulted in the student acquiring the productive knowledge of meaning for one targeted word (i.e., scaffold). This word was spoken 37 times by the lecturer, the most frequently spoken among the 10 targeted words. It is not surprising to have found the student to have only acquired the productive knowledge of meaning for only one of



the ten targeted words as productive knowledge is more difficult to acquire than receptive knowledge (Nation, 2013a). This is because word details must be attended to when a person speaks or writes during language production (Webb and Nation, 2017). As there was no opportunity for the student to use these words productively during the study, having acquired little productive knowledge of vocabulary comes as no surprise. The implications for content teachers is that without a flood of exposure to unknown words, they should not assume that students will have the wherewithal to incidentally acquire the knowledge of unknown words to the point that they can be used productively. As it might not be feasible for a teacher to increase the frequency of exposure to the words through repetition in their lectures, an alternative would be to ensure important words are used when completing coursework. Having students complete tasks that require language output (i.e., speaking and writing) will result in more incidental vocabulary learning than completion of receptive tasks (i.e., listening and speaking) (Webb, 2005).

The effect that frequency has on the learning of vocabulary cannot be considered in isolation. The target word *lexical* was spoken 35 times by the lecturer; however, the assessments did not show the student to have acquired productive knowledge of this word. The word *scaffold* was spoken in six lectures and the word *lexical* was spoken in four lectures. Previous research has shown that the knowledge of words that are more evenly dispersed across language input as being more likely to be incidentally acquired (Reynolds and Ding, 2022). Furthermore, looking at the occurrences of the word *scaffold* show it was spoken by the lecturer most often during lecture 1 ( $n = 11$ ) and lecture 2 ( $n = 19$ ) before being spoken again during lectures 19, 20, 21, and 22. Research has also shown knowledge of new words are better retained if they are first presented in a massed learning condition before being followed up with a spaced learning condition (Uchiyama et al., 2019). This is because the initial massed learning allows language learners to more easily make a form-meaning connection, the initial step necessary for vocabulary learning (Nation, 2013a). Thus, the accumulated effects of frequency, range, and massed learning followed by spaced learning resulted in the productive knowledge of meaning for the word *scaffold* being incidentally acquired by the student. In contrast, the lecturer spoke the word *lexical* for the first time in lecture 17 ( $n = 2$ ) and not in a massed learning condition, potentially reducing the chance for the student to make the form-meaning connection. By lectures 25, 26, and 27 the word *lexical* did occur in a massed learning condition but the study had just ended, eliminating the possibility of a follow up spaced learning condition.

Language learners often perform better at receptive knowledge assessments as they only need to recognize a word passively in language input (Webb and Nation, 2017). This was also found in the incidental vocabulary acquisition post-test results of the student. He showed receptive knowledge of form gains of two words and receptive knowledge of meaning gains of four words. While the difficulty in productive knowledge

assessments can partially explain the difference in the incidental acquisition outcomes between the productive and receptive assessments, it is likely salience was more influential than frequency. This is because the frequency of exposure through the lectures for the acquired words varied considerably. Thus, it is likely that for one reason or another the incidentally acquired words were found to be more salient to the student. Studies involving the incidental acquisition of reading have found as few as three exposures were needed for incidental acquisition to occur if the task was found important to the students (Reynolds, 2015a). Salience can be encouraged in a number of ways but most likely it was due to the student's perception of the importance of the words or the lecturer's emphasis or explanation of the words. Another likely possibility was the context in which the repetitions of the words were very informative, providing an abundance of explication that allowed for the learner to pick up the meaning incidentally (Webb, 2008; Soyoo et al., 2022).

Out of the ten targeted words, eight were explicitly explained by the lecturers and two were not. The words *scaffold* and *lexical* were spoken by the lecturers 37 and 35 times, respectively. The assessment results showed the productive knowledge of meaning of *scaffold* was incidentally acquired by the student, but *lexical* was not. The reason may be that *scaffold* was presented as an educational concept that was first defined and then exemplified. On the contrary, *lexical* was spoken as a part of 11 two-word phrases or chunks rather than a single word, such as *lexical demand*, *lexical coverage*, and *lexical chunk*. Seven out of 11 of these two-word phrases were not explained by the lecturers. Accordingly, the increasing exposure to the word *lexical* may not have increased the student's chance of incidentally acquiring the word as explanations or definitions were not provided for many of these two-word phrases. Hence, even though the exposure through listening to the lectures was relatively frequent, they may not have been considered as recurrences of the same word by the student (Reynolds, 2015b).

The EMI lecturers used the following teacher talk strategies to aid listeners' understanding of the target words: giving examples of use, providing definitions, describing scenarios, providing synonyms, paraphrasing meaning, and highlighting the purpose of their use. Regarding the quality of explanations offered by the lecturers, most of the words were explained with clear definitions and examples, while two words were only spoken in the lectures without explicit definitions or examples. Although definitions and examples were used as supportive strategies for explanations, in some cases, we tend to feel that overly explaining the words might make simple concepts complicated. In particular, the lecturers explained *fallacy* and *recast* by providing demonstrations that were lengthy and potentially construed as difficult to follow, especially for a first-year first-semester English education major. This might have reduced the frequency effect and thereby reducing the likelihood of the student to have incidentally acquired these words. The implication for content teachers is that quality of

contexts and introduction of new vocabulary is important for incidental learning. Teachers are advised to provide high quality definitions, illustrations, and examples when introducing new concepts in their lectures.

Although the student reported several difficulties in listening to the EMI lectures, he also reported the experience as satisfying. He felt this experience helped prepare him for future content learning in the university. This result supports Reynolds and Ding (2022) that reported participants were able to incidentally acquire vocabulary from completion of a task even when the task was found to be difficult. The student felt the difficulties he faced were due to his English proficiency and unfamiliarity with the course content. A similar situation was reported by Ma et al. (2022, p. 96) that found “linguistic and content barriers” prevented Mainland Chinese graduate students studying in Macau from getting the most from their EMI lectures. These are examples of barriers that Peng and Xie (2021) claimed that should be addressed when implementing an EMI policy in schools. It is necessary for schools in Macau to consider what types of support they should provide to students in the period of adjustment as they transition from a non-EMI to an EMI environment. University lecturers should also be aware that many of the students may have inadequate vocabulary knowledge to understand all of their lectures and should be prepared to provide proper scaffolding to ensure effective instruction.

The finding that the student was able to incidentally acquire vocabulary from listening to academic English lectures has given support to previous findings about the use of academic lectures as language input for vocabulary learning (Dang et al., 2022). The student claimed his interest in the lecture topics and their relevance to his future studies and work is what really honed his attention on particular concepts and vocabulary, thereby inducing a state of incidental learning. He also mentioned that the difficulty of understanding particular concepts was reduced by the lecturer’s explanations, thereby highlighting the importance of providing training to lecturers in how to facilitate their students’ incidental vocabulary acquisition using relevant teacher talk strategies (Horst, 2010).

The student’s interview responses highlighted the need for a bridge between secondary and tertiary education. His experience in listening to the EMI lectures underscored the lack of preparation he felt he received from his secondary school English teachers. If given similar opportunities to listen to EMI lectures in secondary school, he thought that would have been a strong motivator for him and his peers to have paid more attention in English class. Likewise, he felt his secondary school English teachers could have used some teaching materials that mimicked those that would be provided in tertiary education instead of only providing some simple reading materials or language exercises. He assumed when formal classroom lecturing occurred in the university that he would be provided opportunities to interact with his teachers

and peers. He had the assumption that these interactions would facilitate the incidental acquisition of vocabulary encountered in class. His assumption has actually been supported in previous research by Hong and Basturkmen (2020) that found most LREs initiated by the students in their study were about field specific vocabulary. Thus, providing such opportunities to students in the EMI courses is highly recommended.

While English for academic purposes courses surely can provide academic support to students, they may not be enough to properly prepare students for EMI lectures. Candid explanations regarding the benefits and potential drawbacks of EMI should be discussed with not only the students but also the lecturers. Teacher development courses that expound Macau’s educational context, the status of the English language, the overview of students’ linguistic knowledge, capacity for EMI learning, and needs, among others, can allow lecturers to better facilitate their students’ content learning. Such training can help maintain quality EMI programs in Macau (Xie, 2021).

Without such open discussions, the student held an unfounded belief that non-native speakers of English would provide EMI lectures that were easier to understand than lectures provided by native speakers of English. He felt non-native speakers of English are better equipped to understand the needs of non-native speaking students. For example, he said they could easily provide on-the-fly first language translations for unknown words. While limited use of the first language in EMI lectures could help students understand the content being delivered (Yu et al., 2020) and quickly provide translations for thorny concepts or vocabulary (Jin and Webb, 2020), this is not an option for lectures that do not share a common language with the students. Thus, he assumed that native speakers, regardless of their training, educational backgrounds, or experiences, would deliver lectures that were more difficult to understand. The student also struggled with his belief that attending lectures delivered in English would help improve his English language but might prevent him from fully understanding the content.

While these are all valid concerns, some of them could have been alleviated through new student orientation. Therefore, not only the lecturers should be provided with professional training but also students should be given proper transitional support and orientation that ensures they fully understand this new educational context. These findings coincide with the existing studies that have found varied forms of student support along with teacher professional development can lead to successful EMI educational outcomes (Chen and Kraklow, 2015; Aizawa and Rose, 2019; Hong and Basturkmen, 2020; inter alia).

## Conclusion

The findings of this case study have implications for Macau higher education. First, the results found a vocabulary size of the most frequent 4,000 word families plus proper

nouns and marginal words allows students to understand 98% of the entire corpus composed of lectures from the *Introduction to English Language Teaching* course. Similar techniques could be applied to other content course subjects besides English language teaching to determine whether mastery of the most frequent 4,000 word family is also adequate for comprehension of EMI course lectures given in Macau's higher education institutions. While modest, the participant was able to incidentally acquire vocabulary knowledge from listening to the lectures. Still, institutions of higher education in Macau cannot assume students without adequate vocabulary knowledge will be able to pick up most of this knowledge incidentally by listening to EMI lectures. As this study's corpus analyses show, some of the lectures presented a lexical burden that might have reduced the likelihood of incidental vocabulary learning. Instead, teacher talk may need to be adjusted, and other pedagogical strategies should be applied to increase the likelihood of incidental learning. To enhance students' content learning and incidental vocabulary acquisition through the listening to EMI lectures, EMI lecturers are advised to use varied strategies while explaining the target concepts and words to students.

The student's responses to the interview questions resulted in seven themes related to the experience of listening to the lectures and the potential for incidental vocabulary gains from this experience. While at times the student found the experience of listening to the EMI lectures difficult, he still felt it was a rewarding experience as it provided him with an opportunity to prepare for the types of lectures he would encounter during formal university studies. He did, however, mention a few unfounded beliefs regarding lectures given by non-native and native speakers of English. This indicated a need for orientation training for incoming first year students. There was also evidence from the interview data that there is a need for bridging the gap between secondary and tertiary EMI education through teacher in-service training.

While this case study was able to underscore the importance of giving due attention to EMI lectures in the Macau educational context, it was not without its limitations. This case study drew on the data collected from a single course in one subject area of higher education. Without corpus analyses having been conducted on other first year courses in other subject areas, the present results are most applicable to the subject of English language teaching at one higher education institution in Macau. Likewise, only one student was recruited to have his incidental learning assessed with vocabulary tests and have his perceptions of the EMI lectures collected through an interview. In addition, the student only listened to the lectures and did not attend the lectures as is the case in formal course enrollment. It is likely that lectures given in a classroom would be attended by students in a different manner, reducing the ecological validity of the current case study. A future comparison study of students attending in-class and out-of-class lectures could be executed to fully assess

the suitability of this method of collected incidental vocabulary acquisition data.

The present study was mainly concerned with the potentials for and the actual incidental acquisition of vocabulary knowledge from listening to EMI lectures. While the use of the three vocabulary assessment measures were able to provide useful information on the student's acquisition of three aspects of target word knowledge from listening to the lectures, these assessments relied on the use of orthographic vocabulary tests of words encountered through aural input. It would have been more suitable to have administered the vocabulary assessments in an aural format rather than a written format (e.g., Ha, 2021, 2022b); however, this route would have also introduced other potential confounding variables (e.g., pronunciation or accent of the language spoken on the test). As research has found a link between aural single-word knowledge and L2 listening comprehension (Cheng et al., 2022), future research should look carefully into the possibility of adapting and validating existing listening vocabulary levels tests for assessing L1-Chinese learners' existing English vocabulary knowledge (e.g., McLean et al., 2015; Ha, 2021, Ha, 2022a). In addition, researchers should consider development of novel incidental vocabulary acquisition tests with spoken language that would be considered as intelligible to the targeted population of students. If such routes are taken, accent and pronunciation should be carefully controlled.

We also would like to make it known that we acknowledge the current debates and the criticisms that have been raised in the vocabulary assessment literature on the use of particular vocabulary assessments reliant on the construct of word family (Stewart et al., 2021; Webb, 2021b); however, we also acknowledge that some of these arguments are supported with empirical findings drawn almost entirely from learners with L1-Japanese backgrounds. Evidence through replicated empirical studies is indeed needed to determine for certain whether word form variation from language input can affect vocabulary acquisition (Reynolds, 2015b; Webb, 2021c).

Likewise, we also acknowledge the inherent limitations of multiple choice or matching formats used for data collection with the standardized and novel assessments in the present study. However, teachers and researchers alike must make a balance between practicality and reliability when assessing vocabulary knowledge (Nation, 2013b). This is exactly what we have done in the present study. Thus, we highly suggest an err on caution when criticizing existing tools at the disposal of vocabulary researchers and instead see whether these criticisms can be supported from multiple L2 learner populations (see Snoder and Laufer, 2022 for an interesting study with L1 Swedish learners of L2 English as an example).

As with all studies with a limited sample from a targeted population, this study's results are likely to have been biased and should not be generalized to the entire first year Macau undergraduate student body. Likewise, only a small number

of words were assessed in this study. Although these words were carefully selected based on their properties and the participant's linguistic background, it is likely that the student could have acquired other vocabulary that did not appear on the assessments. Future studies should address some of these limitations by collecting lectures from multiple courses within the same and different disciplines as well as from different lecturers from different higher education institutions. This could provide more evidence to make more general claims about the Macau higher education context. Likewise, a much larger student sample should be recruited to gather their perceptions regarding higher education EMI lectures and to assess their incidental knowledge of vocabulary of a larger number of items from viewing such lectures. Still, this study does offer valuable insights to future researchers and has shown that this is a valid area of research that deserves further attention.

## Data availability statement

The original contributions presented in this study are included in the article/**Supplementary material**, further inquiries can be directed to the corresponding author.

## Ethics statement

The studies involving human participants were reviewed and approved by Sub-Panel on Social Science and Humanities Research at the University of Macau. Written informed consent to participate in this study was provided by the participants' legal guardian/next of kin.

## Author contributions

BR contributed to conceptualization, methodology, data collection, transcription, resources, supervision, project administration, and funding acquisition, charge of review and editing, and designed the vocabulary assessments. XX and QP conducted formal analysis and was validated by BR. All authors wrote the first draft of the manuscript. QP proofed the

manuscript. XX and QP were in charge of data curation and software. XX was in charge of language translation. All authors contributed to manuscript revision during the review process, read, and approved the submitted version.

## Funding

The research reported in this manuscript was supported by Education Fund of the Macao SAR Government (HSS-UMAC-2021-02).

## Acknowledgments

We appreciate the student's willingness to participate in this study. We also acknowledge the feedback provided by Tom Cobb on the issue of compounds.

## Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

## Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

## Supplementary material

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fpsyg.2022.993445/full#supplementary-material>

## References

- Aizawa, I., and Rose, H. (2019). An analysis of Japan's English as medium of instruction initiatives within higher education: The gap between meso-level policy and micro-level practice. *High. Educ.* 77, 1125–1142. doi: 10.1007/s10734-018-0323-5
- Anthony, L. (2022). *AntWordProfiler (Version 2.0.0) [Computer Software]*. Tokyo: Waseda University.
- Beglar, D. (2010). A rasch-based validation of the vocabulary size test. *Lang. Test.* 27, 101–118. doi: 10.1177/0265532209340194
- Bowles, H., and Murphy, A. C. (2020). *English-Medium Instruction and the Internationalization of Universities*. Cham: Springer International Publishing AG. doi: 10.1080/03057925.2021.1907982



- Brown, D. (2021). Incidental vocabulary learning in a Japanese university L2-English language classroom over a semester. *TESOL J.* 12, 1–17.
- Chen, Y. L. E., and Kraklow, D. (2015). Taiwanese college students' motivation and engagement for English learning in the context of internationalization at home: A comparison of students in EMI and Non-EMI programs. *J. Stud. Int. Educ.* 19, 46–64. doi: 10.1177/1028315314533607
- Cheng, J., Matthews, J., Lange, K., and McLean, S. (2022). Aural single-word and aural phrasal verb knowledge and their relationships to L2 listening comprehension\*. *TESOL Q.* 1–29. doi: 10.1002/tesq.3137 [Epub ahead of print].
- Cobb, T. (2019). *Compleat Web VP*. Available online at: <https://www.lexutor.ca/vp/comp/> (accessed June 15, 2022).
- Dang, T. N. Y. (2020). The potential for learning specialized vocabulary of university lectures and seminars through watching discipline-related TV programs: Insights from medical corpora. *TESOL Q.* 54, 436–459. doi: 10.1002/tesq.552
- Dang, T. N. Y. (2022). Vocabulary in academic lectures. *J. Engl. Acad. Purp.* 58:1101123. doi: 10.1016/j.jeap.2022.101123
- Dang, T. N. Y., Lu, C., and Webb, S. (2022). Incidental learning of single words and collocations through viewing an academic lecture. *Stud. Second Lang. Acquis.* 44, 708–736. doi: 10.1017/S0272263121000474
- Dang, T. N. Y., and Webb, S. (2014). The lexical profile of academic spoken English. *Engl. Specific Purp.* 33, 66–76. doi: 10.1016/j.esp.2013.08.001
- Drisko, J., and Maschi, T. (2015). *Content analysis*. Oxford: Oxford University Press, doi: 10.1093/acprof:oso/9780190215491.001.0001
- Fievez, I., Montero Perez, M., Cornillie, F., and Desmet, P. (2021). Promoting incidental vocabulary learning through watching a French Netflix series with glossed captions. *Comput. Assid. Lang. Learn.* [Epub ahead of print]. doi: 10.1080/09588221.2021.1899244
- Giordano, M. J. (2021). Lexical coverage in dialogue listening. *Lang. Teach. Res.* 136216882198986. doi: 10.1177/1362168821989869 [Epub ahead of print].
- Ha, T. H. (2021). A Rasch-based validation of the Vietnamese version of the listening vocabulary levels test. *Lang. Test. Asia* 11:16. doi: 10.1186/s40468-021-00132-7
- Ha, T. H. (2022b). Vocabulary demands of informal spoken English revisited: What does it take to understand movies, TV programs, and soap operas? *Front. Psychol.* 13:831684. doi: 10.3389/fpsyg.2022.831684
- Ha, T. H. (2022a). Test format and local dependence of items revisited: A case of two vocabulary levels tests. *Front. Psychol.* 12:805450. doi: 10.3389/fpsyg.2021.805450
- Hong, J., and Basturkmen, H. (2020). Incidental attention to academic language during content teaching in two EMI classes in South Korean high schools. *J. Engl. Acad. Purp.* 48:100921. doi: 10.1016/j.jeap.2020.100921
- Hood, M. (2009). "Case study," in *Qualitative Research in Applied Linguistics: A Practical Introduction*, eds J. Heigham and R. Croker (London: Palgrave Macmillan), 66–90.
- Horst, M. (2010). How well does teacher talk support incidental vocabulary acquisition? *Read. Foreign Lang.* 22, 161–180.
- Hu, G., Li, L., and Lei, J. (2014). English-medium instruction at a Chinese university: Rhetoric and reality. *Lang. Policy* 13, 21–40. doi: 10.1007/s10993-013-9298-3
- Jin, Z., and Webb, S. (2020). Incidental vocabulary learning through listening to teacher talk. *Modern Lang. J.* 104, 550–566. doi: 10.1111/modl.12661
- Keuleers, E., and Brysbaert, M. (2010). Wuggy: A multilingual pseudoword generator. *Behav. Res. Methods* 42, 627–633. doi: 10.3758/BRM.42.3.627
- Kobeleva, P. P. (2012). Second language listening and unfamiliar proper names: Comprehension barrier? *REL C J.* 43, 83–98. doi: 10.1177/0033688212440637
- Kung, F.-W. (2013). The more the merrier? Bilingualism in an academic perspective: Exploring the implementation of English-medium instruction in Taiwanese tertiary education. *Asian EFL J.* 15, 8–36.
- Laufer, B. (1998). The development of passive and active vocabulary in a second language: Same or different? *Appl. Linguist.* 19, 255–271. doi: 10.1093/applin/19.2.255
- Leeming, P. (2014). Analysis of the vocabulary size test. *Lang. Educ. J.* 5, 73–87.
- Ma, Y., Yu, S., Reynolds, B. L., and Jiang, L. (2022). A qualitative investigation of Chinese students' willingness to communicate in English in the graduate school EMI classroom. *Engl. Teach. Learn.* 46, 77–98. doi: 10.1007/s42321-021-00087-1
- Macaro, E., and Han, S. (2020). English medium instruction in China's higher education: Teachers' perspectives of competencies, certification and professional development. *J. Multiling. Multicult. Dev.* 41, 219–231. doi: 10.1080/01434632.2019.1611838
- Macau Education, and Youth Affairs Bureau. (2018). *Senior secondary English curriculum guide for schools using English as the primary language of instruction (EMI)*. The Curriculum Development Website. [https://www.dsedj.gov.mo/crdc/guide/data/senior/English1\\_guide\\_e.pdf](https://www.dsedj.gov.mo/crdc/guide/data/senior/English1_guide_e.pdf) (accessed December 20, 2018).
- McLean, S., Kramer, B., and Beglar, D. (2015). The creation and validation of a listening vocabulary levels test. *Lang. Teach. Res.* 19, 741–760. doi: 10.1177/1362168814567889
- Merriam, S. B. (1988). *Case study research in education*. Hoboken, NJ: Jossey-Bass.
- Mohsen, M. A. (2016). The use of computer-based simulation to aid comprehension and incidental vocabulary learning. *J. Educ. Comput. Res.* 54, 863–884. doi: 10.1177/0735633116639954
- Nation, I. S. P. (2006). How large a vocabulary is needed for reading and listening? *Can. Modern Lang. Rev.* 63, 59–82. doi: 10.3138/cmlr.63.1.59
- Nation, I. S. P. (2013a). *Learning vocabulary in another language*, 2nd Edn. Cambridge: Cambridge University Press.
- Nation, I. S. P. (2013b). *What should every EFL teacher know?*. Seoul: Compass Publishing.
- Nation, I. S. P. (2017). *The BNC/COCA Level 6 word family lists (Version 1.0.0) [Data file]*. Available online at: <http://www.victoria.ac.nz/lals/staff/paul-nation.aspx> (accessed on Mar 26, 2022)
- Nation, I. S. P., and Beglar, D. (2007). A vocabulary size test. *Lang. Teach.* 31, 9–13.
- Nation, I. S. P., and Webb, S. (2011). *Researching and analyzing vocabulary*. Boston, MA: Heinle & Heinle.
- Nguyen, C. D., and Boers, F. (2019). The effect of content retelling on vocabulary uptake from a TED talk. *TESOL Q.* 53, 5–29. doi: 10.1002/tesq.441
- Norellie, A.-S., Kestemont, B., Heylen, K., Desmet, P., and Peters, E. (2018). Vocabulary knowledge and listening comprehension at an intermediate level in English and French as foreign languages. An approximate replication study of Stæhr (2009). *Int. J. Appl. Linguist.* 169, 212–231. doi: 10.1075/itl.00013.nor
- Pavia, N., Webb, S., and Faez, F. (2019). Incidental vocabulary learning through listening to songs. *Stud. Second Lang. Acquis.* 41, 745–768. doi: 10.1017/S0272263119000020
- Peng, J. E. (2012). Towards an ecological understanding of willingness to communicate in EFL classrooms in China. *System* 40, 203–213. doi: 10.1016/j.system.2012.02.002
- Peng, J., and Xie, X. (2021). English-medium instruction as a pedagogical strategy for the sustainable development of EFL learners in the Chinese context: A meta-analysis of its effectiveness. *Sustainability* 13:5637. doi: 10.3390/su13105637
- Reynolds, B. L. (2015a). A mixed-methods approach to investigating first- and second-language incidental vocabulary acquisition through the reading of fiction. *Read. Res. Q.* 50, 111–127. doi: 10.1002/rtrq.88
- Reynolds, B. L. (2015b). The effects of word form variation and frequency on second language incidental vocabulary acquisition through reading. *Appl. Linguist. Rev.* 6, 467–497.
- Reynolds, B. L. (2020). The effects of nonce words, frequency, contextual richness, and L2 vocabulary knowledge on the incidental acquisition of vocabulary through reading: More than a replication of Zahar et al. (2001) & Tekmen & Daloglu (2006). *Int. Rev. Appl. Linguist. Lang. Teach.* 58, 75–102. doi: 10.1515/iral-2015-0115
- Reynolds, B. L., and Ding, C. (2022). Effects of word-related factors on first and second language English readers' incidental acquisition of vocabulary through reading of an authentic novel. *Engl. Teach.* 21, 171–191. doi: 10.1108/ETPC-05-2021-0049
- Reynolds, B. L., and Wible, D. (2014). Frequency in incidental vocabulary acquisition research: An undefined concept and some consequences. *TESOL Q.* 48, 843–861. doi: 10.1002/tesq.197
- Reynolds, B. L., and Yu, M. H. (2018). Addressing the language needs of administrative staff in Taiwan's internationalised higher education: Call for an English as a lingua franca curriculum to increase communicative competence and willingness to communicate. *Lang. Educ.* 32, 147–166. doi: 10.1080/09500782.2017.1405017
- Robson, C. (2011). *Real World Research: a Resource for Users of Social Research Methods in Applied Settings*, 3rd Edn. Chichester: Wiley.
- Rodgers, M. P. H., and Webb, S. (2011). Narrow viewing: The vocabulary in related television programs. *TESOL Q.* 45, 689–717. doi: 10.5054/tq.2011.268062

- Rodgers, M. P. H., and Webb, S. (2020). Incidental vocabulary learning through viewing television. *ITL Int. J. Appl. Linguist.* 171, 191–220. doi: 10.1075/itl.18034.rod
- Rott, S. (1999). The effect of exposure frequency on intermediate language learners' incidental vocabulary acquisition and retention through reading. *Stud. Second Lang. Acquis.* 21, 589–619. doi: 10.1017/S0272263199004039
- Smidt, E., and Hegelheimer, V. (2004). Effects of online academic lectures on ESL listening comprehension, incidental vocabulary acquisition, and strategy use. *Comput. Asstd. Lang. Learn.* 17, 517–556. doi: 10.1080/0958822042000319692
- Snoder, P., and Laufer, B. (2022). EFL learners' receptive knowledge of derived words: The case of Swedish adolescents. *TESOL Q.* 1–24. doi: 10.1002/tesq.3101 [Epub ahead of print].
- Soyoo, A., Reynolds, B. L., Shadiev, R., and Vazquez-Calvo, B. (2022). A mixed-methods study of the incidental acquisition of foreign language vocabulary and healthcare knowledge through serious game play. *Comput. Asstd. Lang. Learn.* 1–34. doi: 10.1080/09588221.2021.2021242 [Epub ahead of print].
- Soyoo, A., Reynolds, B. L., Vazquez-Calvo, B., and McLay, K. (2021). Informal digital learning of English (IDLE): A scoping review of what has been done and a look towards what is to come. *Comput. Asstd. Lang. Learn.* doi: 10.1080/09588221.2021.1936562
- Stewart, J., Stoeckel, T., McLean, S., Nation, P., and Pinchbeck, G. G. (2021). What the research shows about written receptive vocabulary testing. *Stud. Second Lang. Acquis.* 43, 462–471. doi: 10.1017/S0272263121000437
- Stæhr, L. S. (2009). Vocabulary knowledge and advanced listening comprehension in English as a foreign language. *Stud. Second Lang. Acquis.* 31, 577–607. doi: 10.1017/S0272263109990039
- Sung, C. C. M. (2020). Investing in English-mediated practices in the EMI university: The case of cross-border mainland Chinese students in Hong Kong. *Lingua* 243:102919. doi: 10.1016/j.lingua.2020.102919
- Tam, I. A., and Reynolds, B. L. (2022). The relationship between extramural English engagement and the vocabulary size of L1 Cantonese speakers in Macau. *ITL Int. J. Appl. Linguist.* 1–35. doi: 10.1075/itl.21003.tam [Epub ahead of print].
- Tang, E. (2011). Non-native teacher talk as lexical input in the foreign language classroom. *J. Lang. Teach. Res.* 2, 45–54. doi: 10.4304/jltr.2.1.45-54
- Tauroza, S., and Allison, D. (1990). Speech rates in British English. *Appl. Linguist.* 11, 90–105. doi: 10.1093/applin/11.1.90
- Toh, G. (2020). Challenges in English-medium instruction (EMI) at a Japanese university. *World Engli.* 39, 334–347. doi: 10.1111/weng.12434
- Tri, D. H., and Moskovsky, C. (2019). English-medium instruction in Vietnamese higher education: A road-mapping perspective. *Issues Educ. Res.* 29, 1319–1336.
- Uchihara, T., Webb, S., and Yanagisawa, A. (2019). The effects of repetition on incidental vocabulary learning: A meta-analysis of correlational studies. *Lang. Learn.* 69, 559–599. doi: 10.1111/lang.12343
- Van Patten, B. (2003). *From input to output: A teacher's guide to second language acquisition*. New York, NY: McGraw-Hill.
- van Zeeland, H., and Schmitt, N. (2013). Incidental vocabulary acquisition through L2 listening: A dimensions approach. *System* 41, 609–624.
- Vilkaitė-Lozdienė, L., and Schmitt, N. (2019). "Frequency as a guide for vocabulary usefulness: High-, mid-, and low-frequency words," in *The Routledge Handbook of Vocabulary Studies*, ed. S. Webb (Milton Park: Routledge), 81–96.
- Webb, S. (2005). Receptive and productive vocabulary learning: The effects of reading and writing on word knowledge. *Stud. Second Lang. Acquis.* 27, 33–52. doi: 10.1017/S0272263105050023
- Webb, S. (2007). The effects of repetition on vocabulary knowledge. *Appl. Linguist.* 28, 46–65. doi: 10.1093/applin/aml048
- Webb, S. (2008). The effects of context on incidental vocabulary learning. *Read. Foreign Lang.* 20, 232–245.
- Webb, S. (2021a). Research investigating lexical coverage and lexical profiling: What we know, what we don't know, and what needs to be examined. *Read. Foreign Lang.* 33, 278–293.
- Webb, S. (2021b). A different perspective on the limitations of size and levels tests of written receptive vocabulary knowledge. *Stud. Second Lang. Acquis.* 43, 454–461. doi: 10.1075/S0272263121000449
- Webb, S. (2021c). The lemma dilemma: How should words be operationalized in research and pedagogy? *Stud. Second Lang. Acquis.* 43, 941–949. doi: 10.1017/S0272263121000784
- Webb, S., and Nation, I. S. P. (2017). *How vocabulary is learned*. Oxford: Oxford University Press.
- Webb, S., and Rodgers, M. P. H. (2009a). The lexical coverage of movies. *Appl. Linguist.* 30, 407–427. doi: 10.1093/applin/amp010
- Webb, S., and Rodgers, M. P. H. (2009b). Vocabulary demands of television programs. *Lang. Learn.* 59, 335–366. doi: 10.1111/j.1467-9922.2009.00509.x
- Webb, S., Sasao, Y., and Ballance, O. (2017). The updated vocabulary levels test: Developing and validating two new forms of the VLT. *ITL Int. J. Appl. Linguist.* 168, 33–69. doi: 10.1075/itl.168.1.02web
- Xie, X. (2021). Book review: English-medium instruction and the internationalization of universities by Bowles H, Murphy A. *RELC J.* 1–3. doi: 10.1177/00336882211048609 [Epub ahead of print].
- Yu, S., Wang, Y., Jiang, L., and Wang, B. (2020). Coping with EMI (English as a medium of instruction): Mainland China students' strategies at a university in Macau. *Innov. Educ. Teach. Int.* 58, 462–472. doi: 10.1080/14703297.2020.1784248
- Zhang, P., and Graham, S. (2020). Vocabulary learning through listening: Comparing L2 explanations, teacher codeswitching, contrastive focus-on-form and incidental learning. *Lang. Teach. Res.* 24, 765–784. doi: 10.1177/1362168819829022
- Zhang, X. (2013). The I don't know option in the vocabulary size test. *TESOL Q.* 47, 790–811. doi: 10.1002/tesq.98
- Zhou, S. (2010). Comparing receptive and productive academic vocabulary knowledge of Chinese EFL learners. *Asian Soc. Sci.* 6, 14–19. doi: 10.5539/ass.v6n10p14



## OPEN ACCESS

## EDITED BY

Ehsan Rassaei,  
Majan University College, Oman

## REVIEWED BY

Nina Daskalovska,  
Goce Delcev University,  
North Macedonia  
Moslem Yousefi,  
Bu-Ali Sina University, Iran

## \*CORRESPONDENCE

Barry Lee Reynolds  
barryreynolds@um.edu.mo

## SPECIALTY SECTION

This article was submitted to  
Language Sciences,  
a section of the journal  
Frontiers in Psychology

RECEIVED 01 July 2022

ACCEPTED 09 August 2022

PUBLISHED 06 September 2022

## CITATION

Lei Y and Reynolds BL (2022) Learning  
English vocabulary from word cards: A  
research synthesis.  
*Front. Psychol.* 13:984211.  
doi: 10.3389/fpsyg.2022.984211

## COPYRIGHT

© 2022 Lei and Reynolds. This is an  
open-access article distributed under  
the terms of the [Creative Commons  
Attribution License \(CC BY\)](#). The use,  
distribution or reproduction in other  
forums is permitted, provided the  
original author(s) and the copyright  
owner(s) are credited and that the  
original publication in this journal is  
cited, in accordance with accepted  
academic practice. No use, distribution  
or reproduction is permitted which  
does not comply with these terms.

# Learning English vocabulary from word cards: A research synthesis

Yuanying Lei<sup>1</sup> and Barry Lee Reynolds<sup>1,2\*</sup>

<sup>1</sup>Faculty of Education, University of Macau, Taipa, Macao SAR, China, <sup>2</sup>Centre for Cognitive and Brain Sciences, University of Macau, Taipa, Macao SAR, China

Researchers' interest in the learning of vocabulary from word cards has grown alongside the increasing number of studies published on this topic. While meta-analyses or systematic reviews have been previously performed, the types of word cards investigated, and the number of word card studies analyzed were limited. To address these issues, a research synthesis was conducted to provide an inclusive and comprehensive picture of how the use of word cards by learners results in vocabulary learning. A search of the Web of Science and Scopus databases resulted in 803 potential studies, of which 32 aligned with the inclusion criteria. Coding of these studies based on an extensive coding scheme found most studies assessed receptive vocabulary knowledge more often than productive vocabulary knowledge, and knowledge of vocabulary form and meaning were assessed more often than knowledge of vocabulary use. Results of effect size plots showed that more of the reviewed studies showed larger effects for the use of paper word cards than digital word cards, and for the use of ready-made word cards than self-constructed word cards. Results also indicated more studies showed larger effects for using word cards in an intentional learning condition compared with an incidental learning condition, and for using word cards in a massed learning condition compared with a spaced learning condition. Although a correlation was found between time spent using word cards and vocabulary learning outcomes, this correlation was not statistically significant. Learners that were more proficient in English learned more words from using word cards than those less proficient. These results suggest that future researchers should report learner proficiency, adopt reliable tests to assess vocabulary learning outcomes, compare the effectiveness of ready-made word cards and self-constructed word cards, and investigate the learning of different aspects of word knowledge. Teachers should provide learners guidance in how to use word cards and target word selection for self-construction of word cards. In addition, teachers should encourage learners to create word cards for incidentally encountered unknown words and use massed learning when initially working with these new words before using spaced learning for later retrieval practice.

## KEYWORDS

word cards, flashcards, vocabulary, receptive knowledge, productive knowledge

## Introduction

Vocabulary knowledge is essential in second language (L2) learning (Barkat and Aminafshar, 2015; Reynolds and Shih, 2019). When learning English as a second language, acquiring vocabulary is “more important than mastering other language skills,” such as listening, speaking, reading, and writing (Lukas et al., 2020, p. 305). This is because vocabulary “acts as the foundation for learners to communicate” using the language (Lukas et al., 2020, p. 305). Learning a second language (L2) involves the learning of thousands of words (Laufer and Hulstijn, 2001; Nation, 2013a). In order to understand novels, newspapers, and spoken English, a vocabulary size of “3,000 to 4,000 word families” is needed (Nation, 2013a, p. 14). Researchers, teachers, and learners are interested in knowing the most direct route to learn so many words to be able to use language for these and other purposes.

Learners often engage in different activities and use different strategies to learn vocabulary. Vocabulary-learning activities are often compared to determine which activity is most effective. It is advantageous to learn vocabulary from word cards. For example, Webb et al. (2020, p. 16) suggested that word cards lead to “relatively large gains” in vocabulary knowledge compared to studying word lists. The strength of learning vocabulary from word cards comes from the fact that this activity is focused, efficient, and effective (Nation, 2013a). It is focused because “more attention can easily be paid to unknown words with the use of word cards” (Reynolds et al., 2020, p. 3). It is efficient because a large number of words “can be learned in a short time” using word cards (Nation, 2013a, p. 439). It is effective because word cards can be used for both “receptive and productive learning” (Nation and Webb, 2011, p. 41). Moreover, learners have been shown to prefer learning vocabulary from word cards compared to other vocabulary learning activities (e.g., Kuo and Ho, 2012). Therefore, word cards were chosen as the focus of the synthesis among a variety of vocabulary learning activities available to learners.

Although there is generally a consensus that learning vocabulary from word cards is advantageous, one must acknowledge other variables could enhance or reduce their effectiveness. Previous researchers have indicated that many variables affect vocabulary learning outcomes regardless of the vocabulary learning strategy employed by learners. For most intentional vocabulary learning strategies—including the use of word cards—these include how the strategy is employed (e.g., Uchihara et al., 2019) and language learner proficiency (e.g., Webb et al., 2020). More specifically for word card use, these include aspects of word knowledge (e.g., Nation, 2013a, Ch. 11) and types of word cards (e.g., Chen and Chan, 2019; Reynolds et al., 2020). Furthermore, word cards can only be effective when learners have been trained and understand how to use them (Reynolds et al., 2020). Therefore, these variables should

be taken into consideration to understand whether word cards are effective for vocabulary learning.

It is worthwhile to conduct a synthesis of the word card literature to allow for generalization of the results reported in primary studies. A synthesis can help us to systematically review the word card literature, thereby providing a clearer picture of the overall effectiveness of word cards. Such a result can be useful for teachers, learners, and researchers, as a research synthesis can provide clear implications for research and teaching. Compared to meta-analysis, which “requires strict inclusion criteria for calculating effect sizes (ESs),” a research synthesis allows for “more varieties of relevant studies to be included” (Yang et al., 2021, p. 472). Therefore, this study gives a systematic and comprehensive review on the past research regarding vocabulary learning from word cards using a research synthesis methodology.

The current synthesis of primary empirical studies brings significance to the field of vocabulary learning from word cards for two main reasons. Firstly, there is growing interest in the effects that word cards have on vocabulary learning, evident through the large number of studies published on this topic. With this large body of research, it is not surprising that some existing meta-analyses and syntheses also touch on this topic. For example, Webb et al. (2020) conducted a meta-analysis to examine the effectiveness of many vocabulary learning activities including the use of word cards. As a meta-analysis requires some strict inclusion criteria, many relevant word card studies had to be eliminated. Similarly, several researchers have synthesized the word card literature. Unfortunately, their focus was on synthesizing the literature on one specific type of word cards rather than all types of word cards (Nakata, 2011; Lin and Lin, 2019; Ji and Aziz, 2021). Therefore, the previous meta-analyses and syntheses have not given an exclusive picture of how word cards lead to vocabulary learning. To fill this gap, this study adopts an inclusive synthesis approach to examine how the learners’ use of word cards can lead to vocabulary learning.

Secondly, there are several potential variables that may affect vocabulary learning from word card use. For example, the effect of the use of digital word cards has been compared to paper word cards (e.g., Azabdaftari and Mozaheb, 2012; Chen and Chan, 2019). Some studies asked learners to self-construct word cards (e.g., Reynolds et al., 2020), while other studies provided word cards to learners (e.g., Oberg, 2011). However, it appears in the previous literature that researchers have not considered whether this could influence the effectiveness of word card use. The use of word cards is most often assumed to be an intentional vocabulary learning strategy. However, some researchers have reported to use word cards as an incidental learning strategy as well (e.g., Reynolds et al., 2020). Researchers have not considered whether the use of word cards is suitable for incidental learning. The literature usually suggests that learners use word cards in a spaced learning condition. However, some researchers have suggested learners to use massed learning as a



large number of repeated encounters with the words will occur (Uchihara et al., 2019). Word cards were also reported to have been used for different amounts of time in previous studies (e.g., Webb et al., 2020). The amount of time spent learning from word cards might influence vocabulary learning. Moreover, most of the previous research involved learners at different levels of proficiency (e.g., Tan and Nicholson, 1997; Nakata, 2008). Different levels of language proficiency might result in varied amounts of vocabulary learning from word card use. In this regard, this study extends the discussion of learning L2 vocabulary through the use of word cards and includes potential variables that may affect the reported effects in the published word card literature.

Practically, the findings of this synthesis have the potential to benefit two stakeholders. Firstly, this study provides some suggestions for researchers who have been investigating vocabulary learning from word card use. The results can provide suggestions for a future research trajectory. Secondly, this study has the potential to provide teachers with advice on how they can incorporate the use of word cards into their classroom teaching and skill training for learners.

## Literature review

In this section, we review relevant vocabulary, word card, and theory literature before summarizing existing findings about the variables of interest to the present synthesis. Doing this helps to situate the research questions that follow. The results of this research synthesis builds on the literature that is covered in this section.

### Previous research syntheses on vocabulary learning from word cards

Previous meta-analysts and synthesists have conducted research related to English vocabulary learning activities and examined this field from different perspectives. For example, Webb et al. (2020) conducted a meta-analysis which focused on studies investigating “four types of intentional vocabulary learning activities, including flashcards, word lists, writing and fill-in-the-blanks” (Webb et al., 2020, p. 1). In their meta-analysis of 22 studies, Webb et al. (2020) found that “both flashcards and word lists led to relatively large gains in vocabulary knowledge while writing and fill-in-the-blanks lead to relatively small gains” (Webb et al., 2020, p. 19). However, their meta-analysis only included studies with treatments that lasted up to 1 day, i.e., studies with treatments that lasted longer than 1 day were excluded (Webb et al., 2020). Uchihara et al. (2019) conducted a meta-analysis which focused on the effects of repetition on incidental vocabulary learning. In their meta-analysis of 26 studies, Uchihara et al. (2019, p. 559) found that “there

was a medium effect of repetition on incidental vocabulary learning.” However, their meta-analysis only included studies that adopted within participants design, i.e., studies that adopted between participants design were excluded. Various research designs deserve investigation, as there are an increasing number of empirical studies that have included separate groups with different interventions (Kose and Mede, 2018; Reynolds and Shih, 2019; Wulandari and Musfiroh, 2020). Both Webb et al.’s (2020) and Uchihara et al.’s (2019) meta-analyses focused on the form and meaning aspects of word knowledge. Other aspects of word knowledge should also be given attention by meta-analysts and synthesists. Other aspects of word knowledge require more rigid attention in vocabulary learning from word cards research (Uchihara et al., 2019), as vocabulary learning involves more than “associating the new words with their meaning” (Nakata, 2011, p. 20).

Nakata (2011) conducted a systematic review on digital word card programs for vocabulary learning. In this systematic review of 9 digital word card programs, Nakata (2011, p. 17) found that most digital word card programs “have been developed in a way that maximize vocabulary learning.” Lin and Lin (2019) conducted a systematic review and meta-analysis on vocabulary learning from digital word card use. In their systematic review and meta-analysis of 33 studies, Lin and Lin (2019) found that there was a positive and large effect of engagement in activities using digital word cards on vocabulary learning. Later, Ji and Aziz (2021) also conducted a systematic review of vocabulary learning from digital word card use. In their systematic review of 18 studies, Ji and Aziz (2021) also found that the use of digital word cards enhanced learners’ vocabulary knowledge. These previous syntheses and meta-analyses gave insights on the effects of digital word card use but did not report on paper word card use or compare digital word cards to paper word cards. It is necessary to synthesize the studies that used digital word cards and paper word cards as it is important to see which type of word cards can result in better vocabulary learning outcomes.

Although previous syntheses have been investigating English language learning activities, few comprehensive syntheses have been conducted that focus on the use of word cards for vocabulary learning. Studies that utilized different research designs and assessed different aspects of word knowledge should be included for analysis, as the existing word card research was implemented in various research designs and assessed various aspects of word knowledge. In addition, various variables that might affect the vocabulary learning from word cards should be extracted from the studies for analysis.

It is evident that there is a growing interest in the effects that word cards have on vocabulary learning. This is shown from the number of different syntheses and meta-analyses that have been conducted on this topic (Nakata, 2011; Elgort, 2017; Lin and Lin, 2019; Kim and Webb, 2022). There is also a growing body of studies on word card use (Chen and Chan, 2019; Reynolds and Shih, 2019; Reynolds et al., 2020). However, the syntheses

and the meta-analyses have not been very comprehensive in terms of the aspects of word knowledge assessed and the types of word cards used. The current synthesis is an attempt to give a systematic and comprehensive review on the past vocabulary learning research with a focus on word card use, hoping to provide some teaching implications and suggestions for future research.

## Theoretical perspectives of vocabulary learning from word cards

There are several theoretical perspectives that have been used to frame previous studies. However, the majority of studies have used the Involvement Load Hypothesis (Laufer and Hulstijn, 2001), the Pimsleur's Memory Schedule (Pimsleur, 1967), or the Dual-Coding Theory (Paivio, 1979). The word card studies included in the current synthesis relied on these theories for their research designs and interpretations of their results.

### Involvement load hypothesis

The use of word cards is regarded as a task that has high involvement. The Involvement Load Hypothesis (ILH) is a "task-induced involvement" theory that consists of "three motivational and cognitive dimensions," i.e., need, search, and evaluation (Laufer and Hulstijn, 2001, p. 2). Need is the "motivational, non-cognitive dimension of involvement" and refers to "whether unknown words are needed to complete a task" (Laufer and Hulstijn, 2001, p. 14; Yanagisawa and Webb, 2021, p. 489). Need is absent when an unknown word is not required (need is 0) (Yanagisawa and Webb, 2021). Need is moderate when it is "imposed by an external agent" (e.g., the learners are required to create word cards for teacher selected words) (need is 1), and it is strong when it is "imposed by the learners themselves" (e.g., the learners wish to create word cards for the incidentally encountered unknown words) (need is 2) (Laufer and Hulstijn, 2001, p. 14; Reynolds et al., 2020; Yanagisawa and Webb, 2021, p. 489). Search and evaluation are the "two cognitive dimensions of involvement" (Laufer and Hulstijn, 2001, p. 14). Search refers to the attempt to find an unknown L2 word's form or its meaning (Laufer and Hulstijn, 2001). Search is absent when the L2 word's form and its meaning are provided in a task (e.g., a reading comprehension task where new words are glossed) (search is 0) (Laufer and Hulstijn, 2001; Yanagisawa and Webb, 2021). Search is moderate when the learners need to find an unknown L2 word's form or its meaning using external resources (e.g., dictionaries or teachers) (search is 1), and it is strong when the learners need to engage in both receptive learning and productive learning (e.g., looking at the L2 word forms and trying to recall the L1 translations, and looking at the L1 translations and trying to recall the L2 word forms on word cards) (search is 2)

(Reynolds et al., 2020; Yanagisawa and Webb, 2021). Evaluation involves "the comparison of a given word with other words" (Laufer and Hulstijn, 2001, p. 14). Evaluation is absent when the learners do not need to decide which word to use (evaluation is 0) (Yanagisawa and Webb, 2021). Evaluation is moderate when it entails recognizing differences between words with a context provided (e.g., a fill-in-the-blanks task with given words) (evaluation is 1), and it is strong when a word must be used in an authentic context (evaluation is 2) (e.g., a composition writing task using target words) (Laufer and Hulstijn, 2001; Yanagisawa and Webb, 2021). The strength of the involvement load can occur in any combination. ILH predicts that "higher involvement in a word induced by the task will result in better retention" (Laufer and Hulstijn, 2001, p. 20).

Some researchers used the ILH as a framework for designing their studies. For studies that used word cards, the involvement load was calculated as 6 out of a possible 6. For example, in Reynolds et al.'s (2020, p. 5) study, learners were required to "construct word cards for unknown words encountered while reading a class textbook." In Reynolds et al.'s (2020, pp. 5–6) study, need was 2 (as the learners initiated the need to understand "the unknown words incidentally encountered during reading class texts"), search was 2 (as word cards were used for both receptive learning, i.e., the learners recalled the L1 translations by looking at the L2 word forms, and productive learning, i.e., the learners recalled the L2 word forms by looking at the L1 translations), and evaluation was 2 (as the learners compared "multiple meanings of the words" and used the chosen word to write a sentence on the word card). However, the ILH can only suggest the predictability of a task being useful or not for vocabulary learning. To address the issue of how memory works in the learning of vocabulary from word cards, the Pimsleur's Memory Schedule (Pimsleur, 1967) is more suitable.

### Pimsleur's memory schedule

Previous researchers have suggested that the traditional way of memorizing words lacks scheduled repetition, which would lead to forgetting (Mondria and Mondria-De Vries, 1994). Repetition is "essential for vocabulary learning" in a foreign language (Nation, 2013a, p. 451). Pimsleur (1967) recommended "a memory schedule" which can be regarded as a guide "for determining the length of time that should occur between repetitions" (Kose and Mede, 2018, p. 5). Teachers can follow this schedule to space the recall of words previously learned by students. In this schedule, the rationale for determining the amount of time before recalling previously learned words is that most of the forgetting occurs after the initial learning of a word (Kose and Mede, 2018). This forgetting will slow down as time passes by if the words are periodically encountered (Kose and Mede, 2018). Pimsleur (1967) suggested how often new words should be repeated in order to keep them in a person's memory. It should be "5 s, 25 s ( $5^2 = 25$  s), 2 min ( $5^3 = 125$  s), 10 min

( $5^4 = 625$  s), and so on” (Nation, 2013a, p. 454). If learners are provided with opportunities for repetition of new words at the right time, their memories will be refreshed, and the retrieval of the words can improve retention.

Some researchers used the Pimsleur’s Memory Schedule (PMS) (Pimsleur, 1967) as a framework for designing their studies. For example, Kose and Mede (2018) investigated the effects of vocabulary learning using digital word cards with a spaced repetition system following the PMS. By enabling learners to repeatedly be exposed to the target words at the right time, the learners in their study demonstrated a high level of vocabulary acquisition. This is because after the initial learning of a target word, the forgetting is very fast, but the forgetting on the second repetition will be slower (Nation, 2013a). Knowledge of vocabulary decreases less rapidly after each repetition of target words if the spacing has been increased (Mondria and Mondria-De Vries, 1994). However, most of the included studies used increased spacing rather than strictly following the PMS. However, if studies do not strictly use words cards only with printed text and instead opt for word cards containing pictorial elements, then the Dual-Coding Theory (Paivio, 1979) should be considered to understand how this added multimedia element affects learning.

### Dual-coding theory

It is possible for the use of word cards to “combine visual and verbal information” to optimize “memorization of words” (Lavoie, 2016, p. 22). The Dual-Coding Theory (DCT) (Paivio, 1979) proposed that cognition occurs in two distinct codes, i.e., a verbal code for language, and a non-verbal code for mental imagery (Sadoski, 2006). When information is processed through two channels (verbal and non-verbal) instead of one, learners can “benefit from an additional or compensatory scaffold that supports L2 vocabulary learning” (Wong and Samudra, 2019, p. 1187). Visual representations of word meanings, such as pictures or multimedia, play an important role in vocabulary learning. In addition, written word forms must also be processed visually and learned as visible units (Sadoski, 2006). Therefore, dual coding of word cards might enhance memory recall of vocabulary.

Previous researchers have used the DCT as a framework for designing their studies. For example, in Lavoie’s (2016) study, the experimental group that used word cards was compared to a control group. The word cards were presented with words and pictures to ensure the verbal and non-verbal information was processed at the same time. The results showed learners progressing in the learning of new words, demonstrating the additive effects of the two sources of input on vocabulary learning (Lavoie, 2016).

## Aspects of word knowledge

Vocabulary learning is not all or nothing. There are different aspects of word knowledge. At the most general level, vocabulary knowledge can be divided into three main categories, i.e., “form, meaning, and use” (Nation, 2013a, p. 48). Form refers to the “spoken form, written form and word parts”; meaning refers to “the connection between form and meaning, concepts, references and associations of a word”; and use refers to the “grammatical functions, collocations and constraints on use of a word” (Nation, 2013a, p. 539). Each of these aspects of word knowledge can be assessed productively or receptively. Receptive and productive vocabulary knowledge refers to the “learning direction” of vocabulary (Nation, 2013a, pp. 51–52). Productive knowledge of a word is what a learner “needs to know in order to use the word while speaking or writing,” while receptive knowledge is what a learner “needs to know to understand a word while reading or listening” (Crow, 1986, p. 242).

Table 1 (Nation, 2013a) lists these aspects of word knowledge, indicating which ones are well dealt with by learning form word cards, and which ones are partly dealt with by this strategy. Ideal learning occurs when vocabulary has been acquired both receptively and productively. Word cards “can be used for both receptive and productive learning” (Nation, 2013a, p. 441). For example, if the learners are using bilingual word cards with the “L1 on one side and the L2 on the other,” “looking at the L1 and trying to recall the L2 form” involves productive knowledge of form (Nation, 2013a, p. 446; Reynolds et al., 2020, p. 5). If the learners are “looking at the L2 and trying to recall the L1 meaning with the word cards,” it involves receptive knowledge of meaning (Reynolds et al., 2020, p. 5).

## Variables that affect learning from word cards

Several variables have the potential of moderating the effectiveness of learning vocabulary from using word cards. These include the type of word cards used (i.e., paper or digital, ready-made or self-constructed), word cards used in different learning conditions (i.e., incidental or intentional, spaced or massed), period of time they are used, or if they are used by learners with different language proficiencies.

### Paper and digital word cards

Paper word cards are defined as word cards made from paper-based materials (Nation, 2013a). The emergence of digital word cards allows learners to learn vocabulary on computers (Nakata, 2008) or mobile devices (Lai et al., 2020). The use of digital word cards can arouse learners’ interest in vocabulary learning (Lin and Lin, 2019) and potentially lead to learning

TABLE 1 Aspects of word knowledge dealt with by learning from word cards (Nation, 2013a, p. 442).

Form	Receptive	What does the word sound like?	✓
		What does the word look like?	✓✓
Meaning	Productive	How is the word written and spelled?	✓✓
		What meaning does this word form signal?	✓✓
	Receptive	What is included in the concept?	✓
		What word form can be used to express this meaning?	✓✓
Use	Receptive	In what patterns does the word occur?	✓
		What words or types of words must we use with this one?	✓
	Productive	In what patterns must we use this word?	✓
		What words or types of words must we use with this one?	✓

✓✓ = well dealt with, ✓ = partly dealt with.

gains (Başoglu and Akdemir, 2010; Azabdaftari and Mozaheb, 2012; Tsai, 2018; Chen and Chan, 2019; Xodabande et al., 2021).

Ready-made and self-constructed word cards

Ready-made word cards, which are prepared by teachers or bought in stores, are common in the language learning classroom. For example, McDonald and Reynolds (2021) presented ready-made cards based on words taken from storybooks for learners. In addition to using ready-made word cards, learners can also acquire vocabulary by self-constructing their own word cards. For example, Reynolds et al. (2020) required learners to construct 10 word cards for each of the 10 readings in a textbook. Previous researchers have indicated that learners might have a strong affective bond with self-constructed word cards (Mondria and Mondria-De Vries, 1994). It is meaningful to know whether a learner should use self-constructed word cards or ready-made word cards. As learners may select the words by themselves for self-constructed word cards, this selection might affect their vocabulary learning.

Intentional and incidental learning conditions

The two broad approaches to vocabulary learning are intentional and incidental. Intentional vocabulary learning can be defined by whether learners know that “they will be tested on their vocabulary learning” (Webb et al., 2020, p. 2). If learners know of an “upcoming vocabulary test,” they may “pay special attention to vocabulary and engage in intentional learning” (Uchihara et al., 2019, p. 561). Incidental vocabulary learning is defined as “the learning that emerges through a meaning-focused comprehension task in which learners are not told of an upcoming vocabulary test” (Uchihara et al., 2019, p. 561). Thus, learners’ awareness of a future assessment differentiates between incidental learning, where learners are “unaware of a subsequent vocabulary test,” and intentional learning, where “they know they will be tested” (Webb et al., 2020, p. 2).

Spaced and massed learning conditions

A massed learning condition refers to a learning condition in which words are repeated “during a single and continuous period of time,” while a spaced learning condition refers to a learning condition in which words are repeated “across a period of time at ever-increasing intervals” (Kose and Mede, 2018, p. 4). Spacing has often been operationalized “within a strictly controlled laboratory setting” in which learners study individual L2 words at different time intervals (Uchihara et al., 2019, p. 574). In this synthesis, the massed learning condition was operationalized as use of word cards within a single day, while the spaced learning condition was operationalized as use of word cards that lasted for more than 1 day (Uchihara et al., 2019). Previous researchers have examined the effect of spacing on vocabulary development. For example, Kuo and Ho (2012, p. 36) found a larger but non-significant effect on vocabulary learning when word cards were used in spaced learning conditions compared to massed learning conditions, because the effects of spaced learning might be reduced by retrieval activities in both learning conditions.

Time spent learning from word cards

Previous researchers were also interested in the amount of time that learners spent on learning from word cards (Webb et al., 2020). In Webb et al.’s (2020) meta-analysis of vocabulary learning activities, results showed that the number of minutes learners spent per word did not significantly influence vocabulary learning. In the present research synthesis, time spent learning from word cards was operationalized as the number of minutes the learners spent learning vocabulary using the cards.

Proficiency level of learners

The Common European Framework of Reference for Languages (CEFR) (Council of Europe, 2001) is “the most



influential language framework in the field of second language teaching and assessment” (Fleckenstein et al., 2020, p. 2). “It describes foreign language competencies in three broad stages which can be divided into six proficiency levels,” i.e., A1/A2 for basic users, B1/B2 for independent users, and C1/C2 for proficient users (Fleckenstein et al., 2020, p. 2). Previous researchers have indicated that more advanced learners usually acquire more vocabulary than less proficient learners, as greater L2 knowledge should help learners to understand and use language (Webb et al., 2020).

### Testing vocabulary knowledge

In the previous research investigating the effects of word card use on vocabulary learning, researchers have used standardized tests and researcher-constructed tests. These tests have been used to assess different aspects of vocabulary knowledge (i.e., receptive and productive knowledge of form, meaning, and use). In this section, the standardized tests and the researcher-constructed tests used in these previous studies are introduced.

#### Standardized tests

Three main standardized tests have been used in the published literature. These include the Vocabulary Size Test (VST) (Nation and Beglar, 2007), the Updated Vocabulary Levels Test (UVLT) (Webb et al., 2017), and the New General Service Lists Test (NGSLT) (Browne et al., 2013). Table 2 provides example items from these standardized tests.

The VST (Nation and Beglar, 2007) was designed to measure a learner’s overall English receptive vocabulary knowledge. It is one of the most popular tests used to measure vocabulary size. The VST consists of 140 multiple-choice items. It consists of “10 sampled target words from each of the 1,000-level word family” lists up to the 14,000 level extracted from the “100,000,000 token British National Corpus” (Reynolds et al., 2020, p. 4). Answering all items correctly indicates that the test taker knows the most frequent “14,000 word families” of English (Reynolds et al., 2020, pp. 4–5).

The UVLT (Webb et al., 2017) allows one to measure the mastery of vocabulary at different frequency levels. Specifically, the “first 1,000 most frequent words” of English to the “fifth 1,000 most frequent words” of English are assessed (Webb et al., 2017, p. 35). A test taker is presented with 30 questions per level. A test taker that scores “at least 26/30 (87%) has achieved mastery of that level” and might then focus on learning words from the next level (Webb et al., 2017, p. 56). However, the stricter criterion of 29/30 is recommended for mastery of the first three (1,000–3,000 word families) levels as those are commonly accepted as the basis for future vocabulary learning.

TABLE 2 Standardized tests of vocabulary knowledge.

Type of test	Example item
The Vocabulary Size Test (VST) (Nation and Beglar, 2007)	Maintain: Can they maintain it? a. keep it as it is b. make it larger c. get a better one than it d. get it
The Updated Vocabulary Levels Test (UVLT) (Webb et al., 2017)	___ formal and serious manner ___ winner of a sporting event ___ building where valuable objects are shown 1. bull 2. champion 3. dignity 4. hell 5. museum 6. solution
The New General Service Lists Test (NGSLT) (Browne et al., 2013)	Include: We are including it. a. paying b. changing c. adding d. reading

The NGSLT (Browne et al., 2013) is “a diagnostic instrument” designed to assess “written receptive knowledge” of the words on the New General Service List (NGSL) (Stoeckel et al., 2018, p. 5; Xodabande et al., 2021, p. 100). The NGSL is comprised of “2,800 high frequency words” and is designed to “provide maximal coverage of texts for learners of English” (Stoeckel et al., 2018, p. 5). The test is “a multiple-choice test that consists of 5 levels, each assessing knowledge of 20 randomly sampled words from a 560-word frequency based level of the NGSL” (Stoeckel et al., 2018, p. 5). The first level represents the most frequent words, the second level represents slightly less frequent words, and so forth. Answering correctly 16 or 17 items out of 20 indicates mastery of that level (Browne et al., 2013).

#### Researcher-constructed tests

Looking at “how well a particular word is known” is called measuring “depth of knowledge,” while looking at “how many words are known” is called measuring “breath of knowledge” (Nation, 2013a, p. 549). Table 3 (Nation, 2013a, p. 442) lists various aspects of what is involved in “knowing a word” and provides a corresponding example test item that has been used in previous research to assess that particular knowledge aspect.

Previous word card research has assessed both receptive form knowledge and productive form knowledge. Receptive knowledge of form refers to whether a learner can recognize the “spoken form of a word, written form of a word, or the parts

TABLE 3 Researcher-constructed tests of vocabulary knowledge.

		Study	Example item
Form	Spoken	Samad and Makingkung (2020)	(P) Read aloud the word and spell it out loud.
	Written	Lukas et al. (2020)	(P) Name the pictures of animals correctly.
	Word parts	N/A	N/A
Meaning	Form and meaning	Fukushima (2019)	(R) Enter the L1 words after the displayed L2 English words. (P) Enter the L2 English words after the displayed L1 words.
	Concept and referents	N/A	N/A
	Associations	Oberg (2011, p. 136)	(R) Many KUT students ____ by bicycle, but some students take the train or bus. A. stick things together B. move around something C. hang out with friends D. insert into a slot E. commute to school F. spend money on G. recognize a face H. study mechanical engineering I. re-charge batteries J. play video games
Use	Grammatical functions	Alhuwaydi (2020)	(P) Specify the part of speech for words.
	Collocations	N/A	N/A
	Constraints on use	N/A	N/A

R, receptive knowledge; P, productive knowledge; N/A, No example item available in the included studies.

in a word” (Nation, 2013a, p. 538). Productive form refers to whether a learner can “pronounce a word correctly, spell and write a word, or produce appropriate inflected and derived forms of a word” (Nation, 2013a, p. 538). For example, Lukas et al. (2020) assessed the productive knowledge of form by having learners complete a word dictation task after they were provided a picture of an animal. Samad and Makingkung (2020) assessed the productive knowledge of form by having learners read aloud a word and spell it out loud.

Previous word card research has also assessed both receptive and productive word meaning. Receptive meaning refers to whether a learner can “recall the appropriate meaning for a word form, understand a range of uses of a word and its central concept, or recall common associations for a word” (Nation, 2013a, p. 538). Productive meaning refers to whether the learner can “produce an appropriate word form to express its meaning, use a word to refer to a range of items, or recall a word when presented with related ideas” (Nation, 2013a, p. 538). For example, Oberg (2011) assessed the receptive knowledge of meaning by having learners take a sentence fill-in-the-blank test (see Table 3). Fukushima (2019) assessed the receptive knowledge of meaning by asking learners to complete a test that required them to provide L1 Japanese for displayed L2 English words. Productive knowledge of meaning was also assessed by asking the learners to provide L2 English words after L1 Japanese words were displayed (Fukushima, 2019).

Previous word card research has assessed both receptive and productive knowledge of use. Receptive use refers to whether a learner can “recognize correct uses of a word in context, recognize appropriate collocations, or tell if a word is a common, formal, or infrequent word” (Nation, 2013a, p. 538). Productive use refers to whether a learner can “use a word in correct grammatical patterns, produce a word with appropriate collocations, or use a word at appropriate times” (Nation, 2013a, p. 538). For example, Alhuwaydi (2020) assessed the productive knowledge of use by asking learners to specify the part of speech for words.

## Purpose of the study

With a view to broadening our understanding of English vocabulary learning from word cards, this study attempts to provide an overview of relevant empirical studies to identify potential variables that may affect learning from word cards. The synthesis was guided by the following research questions (RQs):

RQ 1: What aspects of word knowledge have been investigated in the published word card research literature?

RQ 2: Which type of word cards has a larger effect on vocabulary learning? Digital or paper? Self-constructed or ready-made?

RQ 3: Which condition has a larger effect on vocabulary learning? Incidental or intentional? Spaced or massed?

RQ 4: What is the strength of the correlation between the time spent using word cards and vocabulary learning?

RQ 5: Which language proficiency group (basic, independent, and proficient) can learn the most vocabulary from using word cards?

Methodology

Research synthesis in language learning research

The primary goal of a research synthesis is to integrate empirical research findings by “drawing overall conclusions from many separate investigations that address identical or related hypotheses” (Cooper, 2017, pp. 170–171). The present study is a research synthesis of previous vocabulary learning studies which involve word card activities, attempting to provide generalizations of the practice of using word cards in L2 English learning and the effectiveness of word card usage in L2 English vocabulary development.

A well-designed research synthesis involves seven stages (Cooper, 2017, pp. 32–36): “(1) formulating the problem; (2) searching the literature; (3) gathering information from studies; (4) evaluating the quality of studies; (5) analyzing and integrating the outcomes of studies; (6) interpreting the evidence; and (7) presenting the results.” The seven stages for the current research synthesis are briefly summarized in this section and more detailed explanations are provided in the following sections. The first step is to formulate the problem. In this synthesis, after formulating the five research questions, the key concepts, constructs, and variables were clearly defined to distinguish relevant and irrelevant studies. The second step is searching the literature to identify relevant studies. To locate potential primary studies related to English vocabulary learning from word cards, a comprehensive literature search was conducted in the Web of Science (including SCI-Expanded, SSCI, AHCI, CPCI-S, CPCI-SSH, BKCI-S, BKCI-SSH, ESCI) and Scopus databases using search terms related to vocabulary and word cards. The third step is to gather information from studies. To identify the studies to include in the present research synthesis, a set of inclusion and exclusion criteria was applied to screen the retrieved studies after the literature search. Then, after applying the inclusion and exclusion criteria, a systematic coding process was applied using a coding scheme that helped identify important data for analysis. The reliability of this coding process was checked before further data analysis. The fourth

TABLE 4 The key terms used in the database searches.

Search Terms for Vocabulary	*word* OR vocab* OR collocation* OR “n gram*” OR idiom* OR lexic* OR lexeme* OR “lexical bundle*” OR chunk* OR phras* OR pattern* OR formulaic* OR figurative* OR fixed-frame* OR binomial*
	AND
Search Terms for Word Cards	wordcard* OR flashcard* OR “word card” OR “word cards” OR card OR cards

step is to evaluate the quality of studies. In this synthesis, the Study Design and Implementation Assessment Device (Study DIAD) was used to evaluate the studies (Cooper, 2017). The fifth step is to analyze and integrate the outcomes of the primary studies. Results from the primary word card vocabulary learning studies were combined, identifying systematic data patterns regarding the practice of using word cards and its effects on vocabulary learning development (see Section Results). The sixth and seventh steps are to present and interpret the results. This was done through a discussion of this synthesis (see Section Discussion).

Literature search

To locate potentially relevant studies on English vocabulary learning from word cards, the following electronic databases were comprehensively searched: Scopus and Web of Science (WOS) (including Science Citation Index Expanded, Social Sciences Citation Index, Arts and Humanities Citation Index, Conference Proceedings Citation Index—Science, Conference Proceedings Citation Index—Social Science and Humanities, Book Citation Index—Science, Book Citation Index—Social Sciences and Humanities, Emerging Sources Citation Index). The search covered all document types including journal articles, conference papers, and book chapters. The literature search covered the period from 1945 to July 2021. There were no limits on the publication period for the included studies. All studies were searched and screened for inclusion within the databases, in order to be as inclusive as possible.

The key terms related to vocabulary and word cards were searched in the databases by title, yielding 803 results. A set of inclusion and exclusion criteria were applied to determine the studies to be included in the synthesis (see Inclusion and Exclusion Criteria). Duplicated studies were removed. The combinations of search terms and Boolean operators (“AND” or “OR”) used in the database searches are presented in Table 4.

## Inclusion and exclusion criteria

After the initially eligible studies for the synthesis were identified, they were carefully examined based on a set of inclusion and exclusion criteria. Studies that were based on empirical data were included. Besides quantitative empirical studies, qualitative empirical studies were also included in this synthesis. In addition, studies that were written in English were retrieved in this synthesis, as the published studies on international and English-language journals or conferences were generally regarded as quality studies. Studies with the following features were included:

1. The study was based on empirical data.
2. The study was written in English.

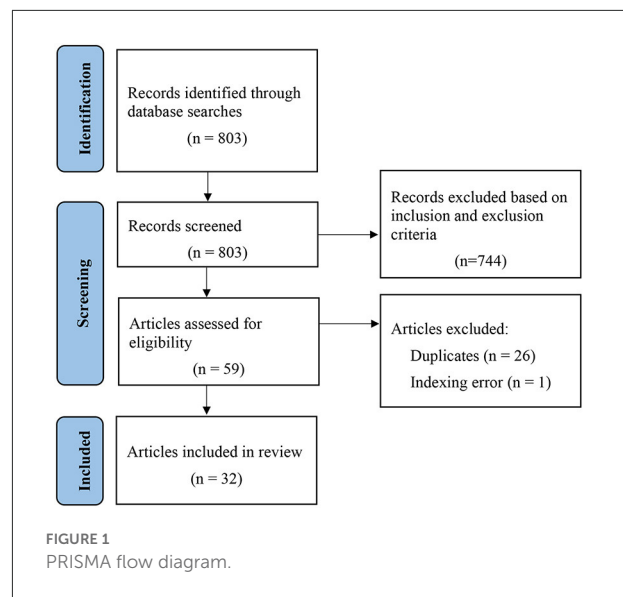
Studies with the following features were excluded:

1. The study was not related to language learning.
2. The study measured non-English language outcomes.
3. The study was not related to vocabulary.
4. The study participants were non-mainstream learners.
5. The study participants were native English speakers.

The 803 studies potentially eligible for the synthesis were then reviewed carefully to identify relevant studies based on the inclusion and exclusion criteria. The titles, abstracts, and full texts (when necessary) of all retrieved papers were reviewed. Seven hundred and forty four were excluded based on the inclusion and exclusion criteria, 26 duplicates were excluded, and 1 was removed due to an indexing error. Overall, the search yielded a sample of 32 studies that were included in the research synthesis. The corresponding full-text documents were obtained. The database search process is presented in a Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) flow diagram (Page et al., 2021) in Figure 1.

## Coding

While the content of every research synthesis coding guide will be unique to the research questions asked, there are certain broad types of information that every synthesist will want to gather from primary research reports (Cooper, 2017). The information to include on a coding guide is classified into eight categories: “(1) the report; (2) the predictor or independent variables; (3) the setting in which the study took place; (4) participant and sample characteristics; (5) the dependent or outcome variables and how they were measured; (6) the type of research design; (7) statistical outcomes and effect sizes; and (8) coder and coding process characteristics” (Cooper, 2017, p. 120). In addition, previous research syntheses on vocabulary learning from word cards coded various variables, including proficiency level and educational level of learners (Webb et al., 2020), number of target words (Wright and Cervetti, 2017; Webb et al., 2020), test timing (Webb et al., 2020) and spacing



(Uchihara et al., 2019). Therefore, after the 32 studies that met the inclusion criteria were identified, Cooper's (2017) coding suggestions and the coded variables of previous research syntheses involving word cards were used to develop the coding scheme for this synthesis.

Specifically, this resulted in seven coding categories: (a) bibliographic information (e.g., author, year of publication), (b) learner characteristics (e.g., sample size, proficiency level), (c) word card characteristics (e.g., origin of word cards, digital integration), (d) methodological characteristics (e.g., study design, theoretical perspective), (e) learning conditions (e.g., spacing), (f) aspects of word knowledge (e.g., receptive form, productive form), and (g) results (e.g., mean of experimental group posttest scores). Supplementary Table 1 provides a detailed description of the coding scheme.

Some of the data was not available in the retrieved research, so the authors of the studies were contacted to request this information. Additional information was gratefully received from four authors (Kose and Mede, 2018; Alhuwaydi, 2020; Hidayat and Yulianti, 2020; Xodabande et al., 2021).

To establish the reliability of the coding procedures, 5 studies (15.63%) were randomly selected and independently coded by a researcher familiar with the process of a research synthesis. Following Boulton and Cobb's (2017) approach, the inter rater reliability was assessed by counting the number of discrepancies between the two researchers' coding. The agreement was found to be 75%. Then, a discussion about the discrepancies was conducted with the researcher. Another 5 studies (15.63%) were randomly selected from the remaining 27 studies and independently coded, then the agreement was found to be 93%. Any remaining disagreements were satisfactorily resolved through discussion, and the coding book was refined where necessary.



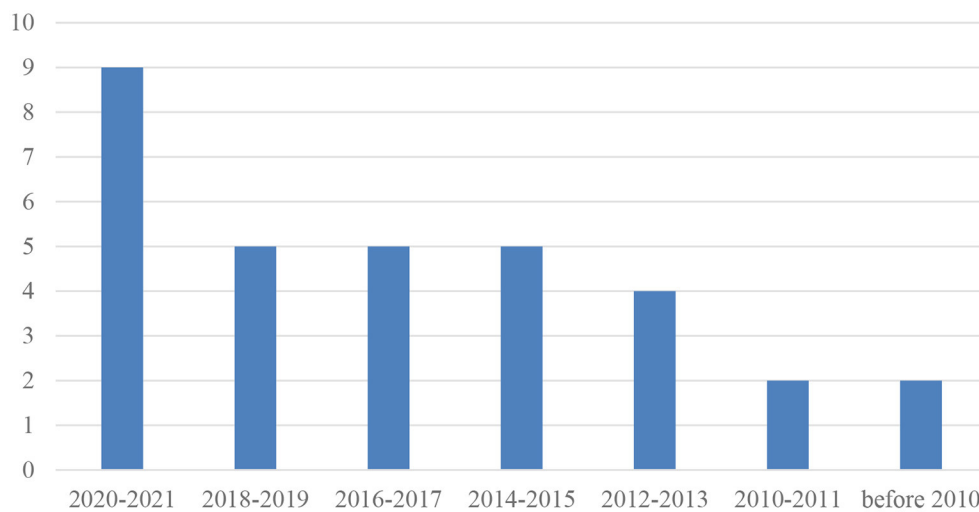


FIGURE 2  
Frequency of year of publication.

## Evaluation of included studies

The Study Design and Implementation Assessment Device (Study DIAD) (Cooper, 2017) was applied to evaluate the included studies. Studies were evaluated with the following four global questions in the Study DIAD (Cooper, 2017, pp. 170–171):

1. Fit Between Concepts and Operations: Were the participants in the study treated and the outcomes measured in a way that is consistent with the definition of the intervention and its proposed effects?
2. Clarity of Casual Inference: Did the research design permit an unambiguous conclusion about the intervention's effectiveness?
3. Generality of Findings: Was the intervention tested on participants, settings, outcomes, and occasions representative of its intended beneficiaries?
4. Precision of Outcome Estimation: Could accurate estimates of the intervention's impact be derived from the study report?"

After detailed evaluation, all 32 included studies were verified to be quality studies.

## Results

### Overview of primary studies

This section provides an overview of the 32 primary studies. [Supplementary Table 2](#) presents the detailed information of these studies in chronological order. Bibliographic information, learner characteristics, word card characteristics,

methodological characteristics and learning conditions are explained below.

### Bibliographic information

In terms of the type of publication, 81% of the studies were journal articles ( $k = 26$ ), followed by conference papers ( $k = 4$ ) and a book chapter ( $k = 1$ ). In terms of the country or region, 25% of the studies were from Taiwan ( $k = 8$ ), 25% were from Iran ( $k = 8$ ), followed by Indonesia ( $k = 4$ ), Turkey ( $k = 4$ ), Japan ( $k = 3$ ), Malaysia ( $k = 1$ ), Macau ( $k = 1$ ), Canada ( $k = 1$ ), New Zealand ( $k = 1$ ), and Saudi Arabia ( $k = 1$ ).

The frequency of year of publication is reported in [Figure 2](#). Although the literature search covered the period from 1945 to July 2021, there was only one study (Tan and Nicholson, 1997) from the 1990s that met the inclusion criteria for the synthesis. Thirty one out of the 32 studies were published from 2008 to 2021. As shown in [Figure 2](#), 9 studies were published between 2020 and 2021, 5 studies between 2018 and 2019, 4 studies between 2016 and 2017, 6 studies between 2014 and 2015, 4 studies between 2012 and 2013, 2 studies between 2010 and 2011, and 2 studies before 2010.

### Learner characteristics

In terms of the proficiency level, 53% ( $k = 17$ ) of the studies reported the proficiency level of the learners. Among the 32 studies, 25% of the studies involved learners at the B1 level ( $k = 8$ ), 12.5% at the A1 level ( $k = 4$ ), followed by B2 ( $k = 3$ ), A2 ( $k = 1$ ), C1 ( $k = 1$ ), and none at the C2 level ( $k = 0$ ). 46.88%

did not report this data ( $k = 15$ ). In terms of the educational level, 93.75% ( $k = 30$ ) of the studies reported the educational level of the learners. Among the 32 studies, 31.25% of the studies involved learners at the university level ( $k = 10$ ), 28.13% at the secondary level ( $k = 9$ ), 21.88% at the primary level ( $k = 7$ ), and 15.63% at the preprimary level ( $k = 5$ ). 6.25% did not report this data ( $k = 2$ ). With respect to L1 backgrounds, 88% ( $k = 28$ ) of the studies reported this data. Among these 32 studies, 25% recruited L1 Mandarin learners ( $k = 8$ ), followed by Turkish ( $k = 5$ ), Indonesian ( $k = 4$ ), Japanese ( $k = 3$ ), Farsi ( $k = 3$ ), Malaysian ( $k = 1$ ), Cantonese ( $k = 1$ ), French ( $k = 1$ ), Persian ( $k = 1$ ), and Arabic ( $k = 1$ ). 12.5% did not report this data ( $k = 4$ ).

## Word card characteristics

In terms of the origin of word cards, 81.25% of the studies used ready-made word cards ( $k = 26$ ), and 18.75% used self-constructed word cards ( $k = 6$ ). In terms of the digital integration, 40.63% of the studies did not use digital word cards ( $k = 13$ ), 31.25% integrated used a computer program with word cards ( $k = 10$ ), and 28.13% used a mobile app with word cards ( $k = 9$ ). In terms of the semantic relatedness, 21.88% of the studies reported the data ( $k = 7$ ). Among the 32 studies, 18.75% of the studies used semantic clustering of words ( $k = 6$ ), 3.13% used thematic clustering of words ( $k = 1$ ), and 78.13% did not report this data ( $k = 25$ ). With respect to the type of assessed vocabulary, 69% of the studies reported the data. Among the 32 studies, 46.88% assessed specific vocabulary knowledge ( $k = 15$ ), 21.88% assessed general vocabulary knowledge ( $k = 7$ ), and 31.25% did not report this data ( $k = 10$ ). In terms of the type of vocabulary test used in previous studies, 90.63% used researcher-constructed tests ( $k = 29$ ), followed by VST ( $k = 2$ ), VLT ( $k = 1$ ), and NGS LT ( $k = 1$ ).

## Methodological characteristics

In terms of the study design, 71.88% of the studies used an independent-group pretest-posttest design ( $k = 23$ ), 18.75% used a single-group pretest-posttest design ( $k = 6$ ), followed by other designs ( $k = 3$ ). In terms of the theoretical perspectives, only 9.38% of the studies reported the use of any theoretical perspective ( $k = 3$ ) for framing the studies. These perspectives included the Involvement Load Hypothesis ( $k = 1$ ) (Laufer and Hulstijn, 2001), the Pimsleur's Memory Schedule ( $k = 1$ ) (Pimsleur, 1967), and the Dual-Coding Theory ( $k = 1$ ) (Paivio, 1979). With respect to a control group, 34.38% of the studies had a control group ( $k = 11$ ), and 66.63% did not have a control group ( $k = 21$ ). In terms of the pretest use, 87.5% conducted a pretest ( $k = 28$ ) and 12.5% did not ( $k = 4$ ). With respect to the test timing, 100% conducted an immediate posttest ( $k = 32$ ),

25% of the studies conducted a delayed posttest ( $k = 8$ ), and 25% conducted both ( $k = 8$ ).

## Learning conditions

In terms of the approaches, 96.88% of the studies applied intentional learning ( $k = 31$ ), and 6.25% applied incidental learning ( $k = 2$ ). With respect to the spacing, 81.25% of the studies reported the data ( $k = 26$ ). Among the 32 studies, 75% asked learners to apply spaced learning ( $k = 24$ ), while 9.38% had learners apply massed learning ( $k = 3$ ), and 3.13% investigated both conditions ( $k = 1$ ). 18.75% did not report this data ( $k = 6$ ).

## Calculation of effect sizes (ESs)

The studies included in the synthesis were not conducted with identical research designs, i.e., the included studies could be single-group pretest-posttest design or independent-groups pretest-posttest design. Due to the discrepancies in the designs, the guidelines suggested by Morris and DeShon (2002, pp. 107–108) and Navarro (2013, p. 382) were followed to calculate the ESs, as described below.

1. In the single-group pretest-posttest design (formula 1):

$$ES = \frac{\text{Mean}_{\text{post}, E} - \text{Mean}_{\text{pre}, E}}{\sqrt{(\text{SD}_{\text{post}, E}^2 + \text{SD}_{\text{pre}, E}^2)}}$$

2. In the independent-groups pretest-posttest design (formula 2):

$$ES = \frac{\text{Mean}_{\text{post}, E} - \text{Mean}_{\text{pre}, E}}{\text{SD}_{\text{pre}, E}} - \frac{\text{Mean}_{\text{post}, C} - \text{Mean}_{\text{pre}, C}}{\text{SD}_{\text{pre}, C}}$$

3. In the independent-groups pretest-posttest design (formula 3; when mean and standard deviation of pretest scores are not available in the published literature):

$$ES = \frac{\text{Mean}_{\text{post}, E} - \text{Mean}_{\text{post}, C}}{(\text{SD}_{\text{post}, E} + \text{SD}_{\text{post}, C})/2}$$

In the above formulas, post = posttest; pre = pretest; E = experimental group; C = control group; SD = standard deviation. In terms of formula 2 and 3, for independent-groups pretest-posttest studies that did not include a control group (e.g., only included two or more experimental groups), an experimental group, i.e., the least interfering experimental group, was treated as a control group. For example, Barkat and Aminafshar's (2015) study did not include a control group. In their study, learners were assigned into three experimental groups, i.e., paper word cards group, digital word cards group, as well as paper and digital word cards group. In this synthesis,

only the first two experimental groups in their study were analyzed, and the second experimental group was treated as a control group.

It has been recommended by researchers that one study should ideally provide only one ES (Light and Pillemer, 1984). As all included studies conducted an immediate posttest, only immediate posttest scores rather than delayed posttest scores were extracted for ES calculation in each study. It should also be mentioned that the included studies did not assess identical aspects of word knowledge. However, the word knowledge was assessed sequentially in most studies. To calculate the ESs, the data of the first receptive knowledge assessment was extracted from each study to prevent practice effects (i.e., the previous test could affect the subsequent test performance) and gather unified data.

Specifically, 78.13% of the studies ( $k = 25$ ) provided means and standard deviations needed for the computation of effect sizes in this synthesis. Conservative estimates of ESs were filled in for the remaining studies ( $k = 7$ ) that had missing data, i.e., assigning ESs of zero, as minimum treatment effect was assumed (Light and Pillemer, 1984).

When the ESs were calculated, an effect direction plot, i.e., a visual display of non-standardized effects across included studies, was then generated (Thomson and Thomas, 2013). In addition, an effect size plot was constructed, i.e., ESs were categorized by their size and visually presented. This synthesis method was utilized to answer RQ 2, 3, and 5, which, respectively, concern the type of word cards used, word cards used in different learning conditions, and word cards used by learners with different language proficiencies.

## Research question 1: What aspects of word knowledge have been investigated in the published word card research literature?

The first research question concerns the aspects of word knowledge investigated in the previous studies. To examine what aspects of word knowledge were investigated by researchers, 29 out of the 32 studies (91%) which indicated the aspect of word knowledge assessed were included for analysis (Tan and Nicholson, 1997; Nakata, 2008; Başoglu and Akdemir, 2010; Oberg, 2011; Azabdaftari and Mozaheb, 2012; Komachali and Khodareza, 2012; Kuo and Ho, 2012; Chien, 2013, 2015; Nikoopour and Kazemi, 2014; Barkat and Aminafshar, 2015; Hamzehbagi and Bonyadi, 2015; Özer and Koçoglu, 2015; Galedari and Basiroo, 2016; Lavoie, 2016; Aminafshar, 2017; Saputri, 2017; Wu et al., 2017; Chen and Chan, 2019; Fukushima, 2019; Reynolds and Shih, 2019; Alhuwaydi, 2020; Hidayat and Yulianti, 2020; Lukas et al., 2020; Samad and Makingkung, 2020; Wulandari and Musfiroh, 2020; Yüksel et al., 2020; Xodabande et al., 2021).

Firstly, the aspects of word knowledge assessed in each of the 29 studies were coded. Secondly, the frequency of studies that assessed each aspect of word knowledge (i.e., how many studies out of the 29 studies investigated the different types of word knowledge) was calculated. Results are presented in Figure 3. In these 29 studies, 72.41% assessed receptive knowledge of meaning (RM) ( $k = 21$ ), 41.38% assessed receptive knowledge of form (RF) ( $k = 12$ ), 27.59% assessed productive knowledge of form (PF) ( $k = 8$ ), 20.69% assessed productive knowledge of meaning (PM) ( $k = 6$ ), 13.79% assessed receptive knowledge of use (RU) ( $k = 4$ ), and 6.90% assessed productive knowledge of use (PU) ( $k = 2$ ).

Based on the frequency of studies that assessed each aspect of word knowledge, most studies assessed receptive vocabulary knowledge more often than productive vocabulary knowledge. In addition, knowledge of vocabulary form and meaning were assessed more often than knowledge of vocabulary use.

## Research question 2: Which type of word cards has a larger effect on vocabulary learning? Digital or paper? Self-constructed or ready-made?

The second research question concerns the types of word cards that were used in the previous studies. To examine the effects of word card type on vocabulary learning, all 32 studies that provided the necessary data were analyzed. The digital integration and the origin of word cards for each of the 32 studies were coded and their ESs were calculated.

Firstly, results concerning the digital integration showed that among the studies, 50% of the studies used paper word cards that did not contain digital integration ( $k = 16$ ), 31.25% used digital word cards in a computer program ( $k = 10$ ), and 28.13% used digital word cards in a mobile app ( $k = 9$ ). Digital word cards (59.38%,  $k = 19$ ) were used more often by researchers than paper word cards (50%,  $k = 16$ ). Specifically, digital word cards in a computer program (31.25%,  $k = 10$ ) were used more often than digital word cards in a mobile app (28.13%,  $k = 9$ ).

Secondly, results concerning the origin of the word cards showed that among the studies, 81.25% used ready-made word cards ( $k = 26$ ), and 18.75% used self-constructed word cards ( $k = 6$ ). Among the two word card types, ready-made word cards were used more often by researchers than self-constructed word cards.

Thirdly, the effect direction plot was constructed and visually presented in Table 5. Arrows were used to indicate reported effect direction (positive effect ▲, negative effect ▼, or no change ◀▶) (Thomson and Thomas, 2013). Arrows were also used to indicate sample size (large arrow=sample size equals or >50, small arrow=sample size smaller than 50) (Thomson and Thomas, 2013). Among the 32 studies, 75% reported a positive effect ( $k = 24$ ) and 25% reported no change ( $k = 8$ ). There

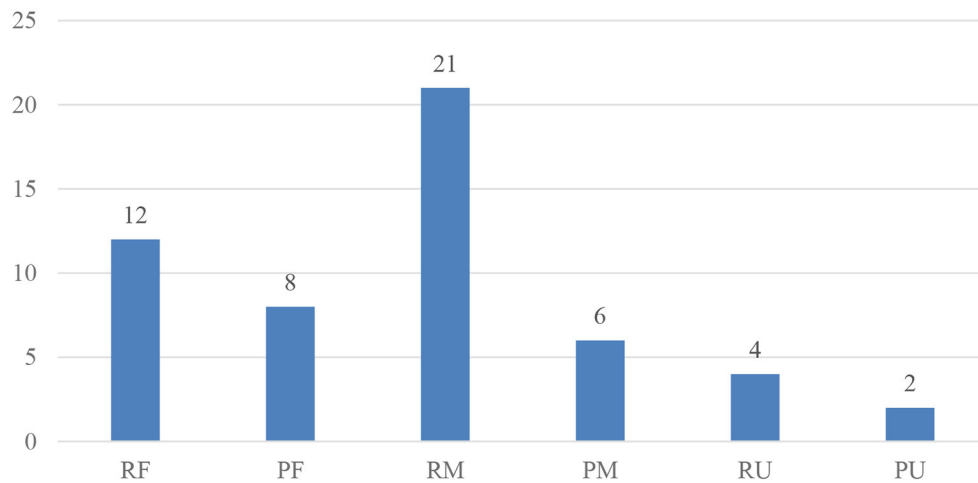


FIGURE 3

Frequency of word knowledge type ( $k = 29$ ). Max = 29 as only 29 of the 32 studies provided necessary data. RF, receptive knowledge of form; PF, productive knowledge of form; RM, receptive knowledge of meaning; PM, productive knowledge of meaning; RU, receptive knowledge of use; PU, productive knowledge of use. As more than one aspect of word knowledge might have been assessed within a single study, the total is higher than 29.

was a trend showing most studies that applied the word cards, either digital or paper, ready-made or self-constructed, showed a positive effect on vocabulary learning.

Lastly, the effect size plot was constructed, i.e., ESs were categorized by their size and visually presented in Figure 4. ESs for these studies were interpreted according to Cohen's guidelines, i.e.,  $< .2$  is negligible,  $.2$  is small,  $.5$  is medium, and  $.8$  is large (Larson-Hall, 2010). In terms of digital integration, studies that used digital word cards showed varied effects, i.e., small (17%), medium (11%) and large effects (33%). Studies that used paper word cards also showed small (17%), medium (11%) and large effects (50%). In terms of the origin of the word cards, studies that used self-constructed word cards showed medium (20%) and large (20%) effects. Studies that used ready-made word cards showed small (15%), medium (11%) and large (44%) effects.

Based on the results of the effect size plot, more of the reviewed studies showed a larger effect for the use of paper word cards compared to digital word cards. In addition, more of the reviewed studies showed a larger effect for the use of ready-made word cards than self-constructed word cards.

### Research question 3: Which condition has a larger effect on vocabulary learning? Incidental or intentional? Spaced or massed?

The third research question concerns the learning conditions that were applied in the previous studies. To

examine the effects of learning conditions on vocabulary learning, all 32 studies that provided the data on the approach (i.e., incidental or intentional learning condition) were analyzed, then 26 studies (81%) that provided the data on spacing (i.e., spaced or massed learning condition) were analyzed (Nakata, 2008; Başoglu and Akdemir, 2010; Azabdaftari and Mozaheb, 2012; Komachali and Khodareza, 2012; Kuo and Ho, 2012; Chien, 2013, 2015; Nikoopour and Kazemi, 2014; Barkat and Aminafshar, 2015; Özer and Koçoglu, 2015; Galedari and Basirloo, 2016; Lavoie, 2016; Aminafshar, 2017; Saputri, 2017; Wu et al., 2017; Kose and Mede, 2018; Tsai, 2018; Chen and Chan, 2019; Fukushima, 2019; Reynolds and Shih, 2019; Hidayat and Yulianti, 2020; Lai et al., 2020; Lukas et al., 2020; Reynolds et al., 2020; Yüksel et al., 2020; Xodabande et al., 2021). The learning approach (i.e., incidental learning or intentional learning) and spacing (i.e., spaced or massed learning condition) for each of the studies was coded and their ESs were calculated.

Firstly, the approach results showed that among the 32 studies, 93.75% of the studies was conducted in an intentional learning condition ( $k = 30$ ), and 9.38% in an incidental learning condition ( $k = 3$ ). An intentional learning condition was used much more often than an incidental learning condition.

Secondly, the spacing results showed that among the 26 studies, 92.31% of the studies used a spaced learning condition ( $k = 24$ ), and 11.54% used a massed learning condition ( $k = 3$ ), and 3.85% used both learning conditions ( $k = 1$ ). More studies used a spaced learning condition than a massed learning condition or both conditions.

Thirdly, the effect direction plot was constructed and visually presented in Table 6. Among the 32 studies, 75% reported a positive effect ( $k = 24$ ) and 25% reported no



TABLE 5 Effect direction plot for type of word cards ( $k = 32$ ).

Studies	Sample size (E/C)	Digital integration	Origin of word cards	Outcomes
Tan and Nicholson (1997)	42	Paper	Ready-made	▲
Nakata (2008)	67/74#	Digital (computer program)	Ready-made	▲
Başoğlu and Akdemir (2010)	29/29	Paper and digital (mobile app)	Ready-made	▲
Oberg (2011)	28/36#	Digital (computer program)	Ready-made	▲
Azabdaftari and Mozaheb (2012)	40/40#	Paper and digital (mobile app)	Ready-made	▲
Komachali and Khodareza (2012)	25/25	Paper	Ready-made	▲
Kuo and Ho (2012)	30/30#	Paper	Ready-made	▲
Chien (2013)	76	Digital (computer program)	Self-constructed	◀▶
Nikoopour and Kazemi (2014)	109	Digital (mobile app)	Ready-made	◀▶
Barkat and Aminafshar (2015)	15/15#	Digital (computer program)	Ready-made	▲
Chien (2015)	64	Digital (computer program)	Self-constructed	◀▶
Hamzehbagi and Bonyadi (2015)	30/30	Paper	Ready-made	▲
Lavoie (2016)	39/15	Paper	Ready-made	▲
Özer and Koçoğlu (2015)	89	Digital (computer program)	Ready-made	◀▶
Galedari and Basiroo (2016)	30/30	Paper	Ready-made	▲
Aminafshar (2017)	15	Digital (computer program)	Ready-made	◀▶
Saputri (2017)	13	Paper	Ready-made	▲
Wu et al. (2017)	10/10	Paper	Ready-made	▲
Kose and Mede (2018)	17/17	Digital (mobile app)	Self-constructed	▲
Tsai (2018)	9/9#	Paper and digital (mobile app)	Ready-made	▲
Chen and Chan (2019)	48/50	Paper and digital (computer program)	Ready-made	▲
Fukushima (2019)	30	Digital (mobile app)	Ready-made	▲
Reynolds and Shih (2019)	100	Paper	Self-constructed	▲
Alhuwaydi (2020)	42	Digital (mobile app)	Self-constructed	▲
Hidayat and Yulianti (2020)	27/26#	Digital (computer program)	Ready-made	◀▶
Lai et al. (2020)	38/20	Digital (mobile app)	Ready-made	▲
Lukas et al. (2020)	52	Paper	Ready-made	▲
Reynolds et al. (2020)	50	Paper	Self-constructed	▲
Samad and Makingkung (2020)	20	Paper	Ready-made	◀▶
Wulandari and Musfiroh (2020)	34/33	Paper	Ready-made	▲
Yüksel et al. (2020)	57	Digital (computer program)	Ready-made	▲
Xodabande et al. (2021)	36/19	Paper and digital (mobile app)	Ready-made	▲

E, experimental group; C, control group. # = an experimental group, i.e., the least interfering experimental group was treated as a control group. ▲ or ▲ = positive effect, ▼ = negative effect, ▶▶ or ◀◀ = no change. Large arrow = sample size equals or > 50, small arrow = sample size smaller than 50.

change ( $k = 8$ ). There was a trend showing most studies that applied these learning conditions showed a positive effect on vocabulary learning.

Lastly, the effect size plot was constructed, i.e., ESs were categorized by their size and visually presented in Figure 5. In terms of the approach, the only two studies that used an incidental learning condition showed negligible effects. Studies that used an intentional learning condition showed small (13%), medium (13%) and large (42%) effects. In terms of spacing, studies that used a spaced learning condition showed small (8%), medium (17%), and large (42%) effects. All studies that used a massed learning condition showed large (100%) effects.

Based on the results of the effect size plot, more of the reviewed studies showed a larger effect for using word cards in

an intentional learning condition compared with an incidental learning condition. In addition, more of the reviewed studies showed a larger effect for using word cards in a massed learning condition compared with a spaced learning condition.

#### Research question 4: What is the strength of the correlation between time spent using word cards and vocabulary learning?

The fourth research question concerns the word card usage time. To examine how time spent using word cards correlates

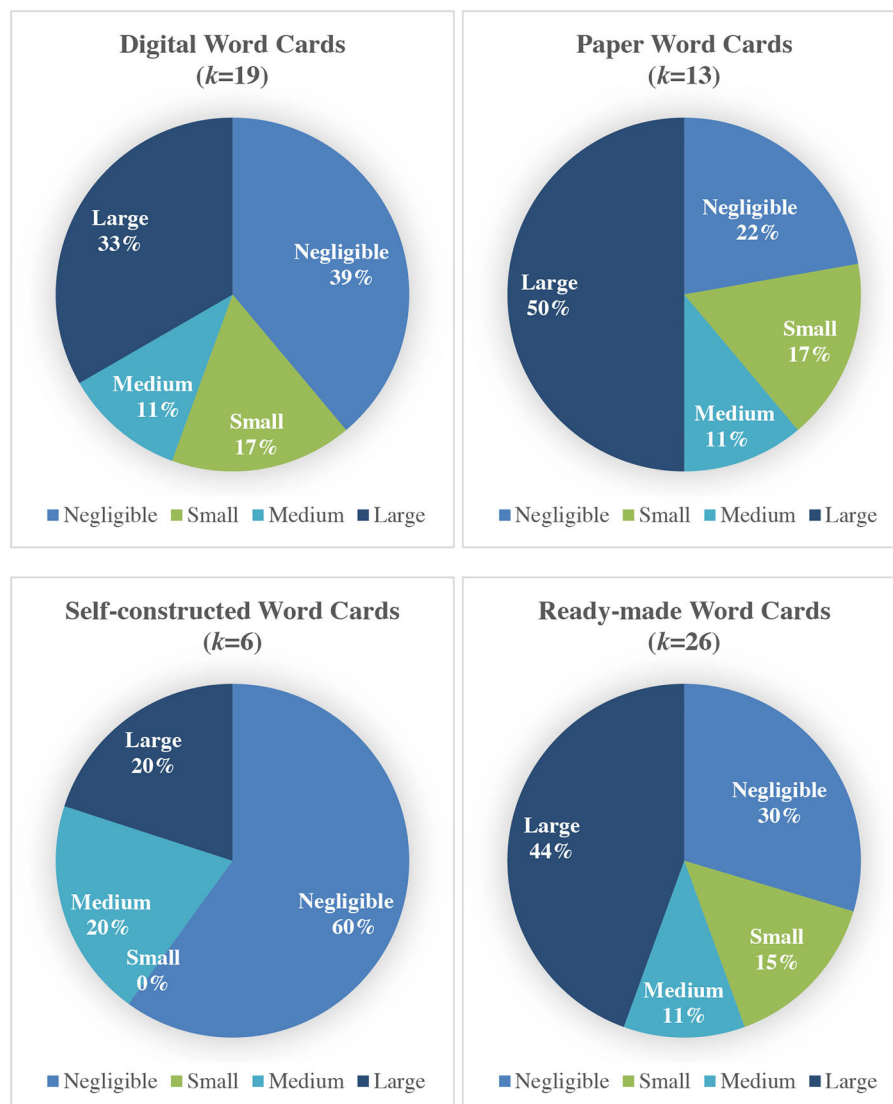


FIGURE 4  
Effect size plot for type of word cards ( $k = 32$ ).

with vocabulary learning, 14 out of the 32 studies (43.75%) that provided the data on time spent using word cards by the learners and the posttest mean scores on vocabulary learning were included for analysis (Tan and Nicholson, 1997; Oberg, 2011; Kuo and Ho, 2012; Galedari and Basirou, 2016; Wu et al., 2017; Kose and Mede, 2018; Tsai, 2018; Chen and Chan, 2019; Fukushima, 2019; Reynolds and Shih, 2019; Hidayat and Yulianti, 2020; Lukas et al., 2020; Reynolds et al., 2020; Xodabande et al., 2021).

Firstly, the time spent using word cards by the learners and the posttest mean scores on vocabulary learning were extracted from the 14 studies. Secondly, the assumptions for Spearman's Rho correlation were checked. The first assumption is that

the data has to be ordinal, interval or ratio, and the second assumption is that the data has to be monotonically related, i.e., one variable increases (or decreases), the other variable also increase (or decreases) (Prion and Haerling, 2014). The extracted data is ratio, i.e., has a true or meaningful zero. In addition, the variables have a monotonic increasing relationship.

Thirdly, a Spearman's Rho correlation was run to determine the relationship between the time spent using word cards and the posttest mean scores on vocabulary learning. Spearman's rho,  $r_s$ , for these studies were interpreted according to the following guidelines: 0 to .2 is negligible, .21 to .4 is weak, .41 to .6 is moderate, .61 to .80 is strong, and .81 to 1 is very strong (Prion and Haerling, 2014). There was a

TABLE 6 Effect direction plot for learning conditions ( $k = 32$ ).

Studies	Sample size (E/C)	Approach	Spacing	Outcomes
Tan and Nicholson (1997)	42	Intentional learning	N/A	▲
Nakata (2008)	67/74#	Intentional learning	Spaced learning	▲
Başoglu and Akdemir (2010)	29/29	Intentional learning	Spaced learning	▲
Oberg (2011)	28/36#	Intentional learning	N/A	▲
Azabdaftari and Mozaheb (2012)	40/40#	Intentional learning	Spaced learning	▲
Komachali and Khodareza (2012)	25/25	Intentional learning	Spaced learning	▲
Kuo and Ho (2012)	30/30#	Intentional learning	Massed and spaced learning	▲
Chien (2013)	76	Intentional learning	Spaced learning	◀▶
Nikoopour and Kazemi (2014)	109	Intentional learning	Spaced learning	◀▶
Barkat and Aminafshar (2015)	15/15#	Intentional learning	Spaced learning	▲
Chien (2015)	64	Intentional learning	Spaced learning	◀▶
Hamzehbagi and Bonyadi (2015)	30/30	Intentional learning	N/A	◀▶
Lavoie (2016)	39/15	Intentional learning	Spaced learning	▲
Özer and Koçoğlu (2015)	89	Incidental learning	Spaced learning	◀▶
Galedari and Basiroo (2016)	30/30	Intentional learning	Spaced learning	▲
Aminafshar (2017)	15	Intentional learning	Spaced learning	◀▶
Saputri (2017)	13	Intentional learning	Spaced learning	▲
Wu et al. (2017)	10/10	Intentional learning	Massed learning	▲
Kose and Mede (2018)	17/17	Intentional learning	Spaced learning	▲
Tsai (2018)	9/9#	Intentional learning	Massed learning	▲
Chen and Chan (2019)	48/50	Intentional learning	Spaced learning	▲
Fukushima (2019)	30	Intentional learning	Spaced learning	▲
Reynolds and Shih (2019)	100	Intentional learning	Spaced learning	▲
Alhuwaydi (2020)	42	Intentional learning	N/A	▲
Hidayat and Yulianti (2020)	27/26#	Intentional learning	Spaced learning	◀▶
Lai et al. (2020)	38/20	Intentional learning	Spaced learning	▲
Lukas et al. (2020)	52	Intentional learning	Spaced learning	▲
Reynolds et al. (2020)	50	Intentional and incidental learning	Spaced learning	▲
Samad and Makingkung (2020)	20	Intentional learning	N/A	◀▶
Wulandari and Musfiroh (2020)	34/33	Intentional learning	N/A	▲
Yüksel et al. (2020)	57	Intentional learning	Spaced learning	▲
Xodabande et al. (2021)	36/19	Intentional learning	Spaced learning	▲

E, experimental group; C, control group. # = an experimental group, i.e., the least interfering experimental group, was treated as a control group. N/A = not available in the publication or from the authors. ▲ or ▼ = positive effect, ◀ or ▶ = negative effect, ◀▶ or ▶◀ = no change. Large arrow = sample size equals or > 50, small arrow = sample size smaller than 50.

positive and weak correlation between the time spent using word cards and the vocabulary learning outcomes ( $r_s = .396$ ,  $p = .161$ , and  $n = 14$ ). The correlation coefficient value of .396 confirmed there was a positive and weak correlation between the two variables, meaning that both variables moved in the same direction. The  $p$ -value of .161 showed that there was not enough evidence to show the correlation was significant.

Based on the results of the Spearman's Rho correlation, a weak relationship was shown between time spent using word cards and vocabulary learning outcomes; however, this relationship was not found to be statistically significant.

## Research question 5: Which language proficiency group (basic, independent, or proficient) can learn the most vocabulary from using word cards?

The fifth research question concerns the proficiency level of learners that were assessed in the previous studies. To examine which proficiency group learned the most vocabulary from using word cards, 17 out of the 32 studies (53.13%) were included for analysis (Tan and Nicholson, 1997; Nakata, 2008; Başoglu and Akdemir, 2010; Oberg, 2011; Azabdaftari and Mozaheb, 2012; Komachali and Khodareza, 2012; Chien, 2013, 2015; Hamzehbagi and Bonyadi, 2015; Özer and Koçoğlu, 2015; Kose

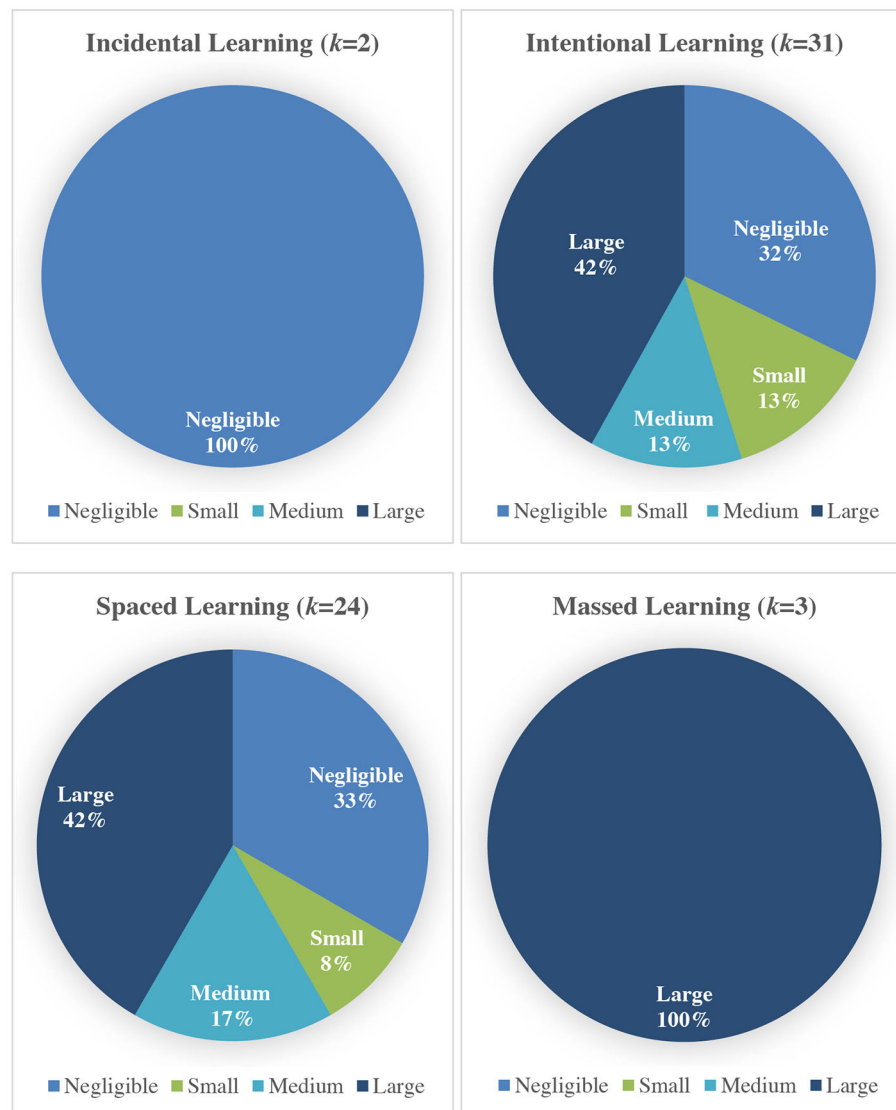


FIGURE 5  
Effect size plot for learning conditions ( $k = 32$ ).

and Mede, 2018; Fukushima, 2019; Reynolds and Shih, 2019; Alhuwaydi, 2020; Hidayat and Yulianti, 2020; Reynolds et al., 2020; Yüksel et al., 2020).

Firstly, the proficiency level in each of the 17 studies was coded and their ESs were calculated. Secondly, the frequency of the studies that assessed proficiency level (i.e., how many studies out of the 17 studies investigated the proficiency level) was calculated. Results showed that among the 17 studies, 47.06% involved learners at the B1 level ( $k = 8$ ), 23.53% at the A1 level ( $k = 4$ ), followed by B2 ( $k = 3$ ), A2 ( $k = 1$ ), C1 ( $k = 1$ ), and none in the C2 ( $k = 0$ ). Learners at the B1 level were involved more often than learners at any other proficiency level.

Thirdly, the effect direction plot was constructed and visually presented in Table 7. Among the 17 studies, 70.59% reported a positive effect ( $k = 12$ ) and 29.41% reported no change ( $k = 5$ ). There was a trend showing most studies that assessed the proficiency level of learners showed a positive effect on vocabulary learning. Lastly, the effect size plot was constructed, i.e., ESs were categorized by their size in terms of basic (i.e., A1 or A2 level), independent (i.e., B1 or B2 level) and proficient (i.e., C1 or C2 level) level group and visually presented in Figure 6. In terms of proficiency level of learners, all studies that assessed basic learners showed negligible effects. Studies that assessed independent learners showed varied effects, i.e., small



TABLE 7 Effect direction plot for proficiency level of learners ( $k = 17$ ).

Studies	Sample size	Proficiency level of learners	Outcomes
(E/C)			
Tan and Nicholson (1997)	42	A1	▲
Nakata (2008)	67/74#	B1	▲
Başoglu and Akdemir (2010)	29/29	B1	▲
Oberg (2011)	28/36#	B2	▲
Azabdaftari and Mozaheb (2012)	40/40#	C1	▲
Komachali and Khodareza (2012)	25/25	B1	▲
Chien (2013)	76	B1	◀▶
Chien (2015)	64	B1	◀▶
Hamzehbagi and Bonyadi (2015)	30/30	A2	◀▶
Özer and Koçoğlu (2015)	89	A1	◀▶
Kose and Mede (2018)	17/17	B1	▲
Fukushima (2019)	30	B1	▲
Reynolds and Shih (2019)	100	B2	▲
Alhuwaydi (2020)	42	B2	▲
Hidayat and Yulianti (2020)	27/26#	A1	◀▶
Reynolds et al. (2020)	50	A1	▲
Yüksel et al. (2020)	57	B1	▲

E, experimental group; C, control group. # = an experimental group, i.e., the least interfering experimental group was treated as a control group. ▲ or ▲ = positive effect, ▼ = negative effect, ◀▶ or ▶◀ = no change. Large arrow = sample size equals or >50, small arrow = sample size smaller than 50.

(20%), medium (20%) and large (30%) effects. The 1 study that assessed proficient learners showed a large (100%) effect.

Based on the results of the effect size plot, learners that were more proficient in English learned more words from using word cards than those less proficient in English. More specifically, learners at the proficient level (i.e., C1 or C2 level) learned the most vocabulary from using word cards, followed by independent (i.e., B1 or B2 level) learners and finally basic (i.e., A1 or A2 level) learners.

## Discussion

### Aspects of word knowledge

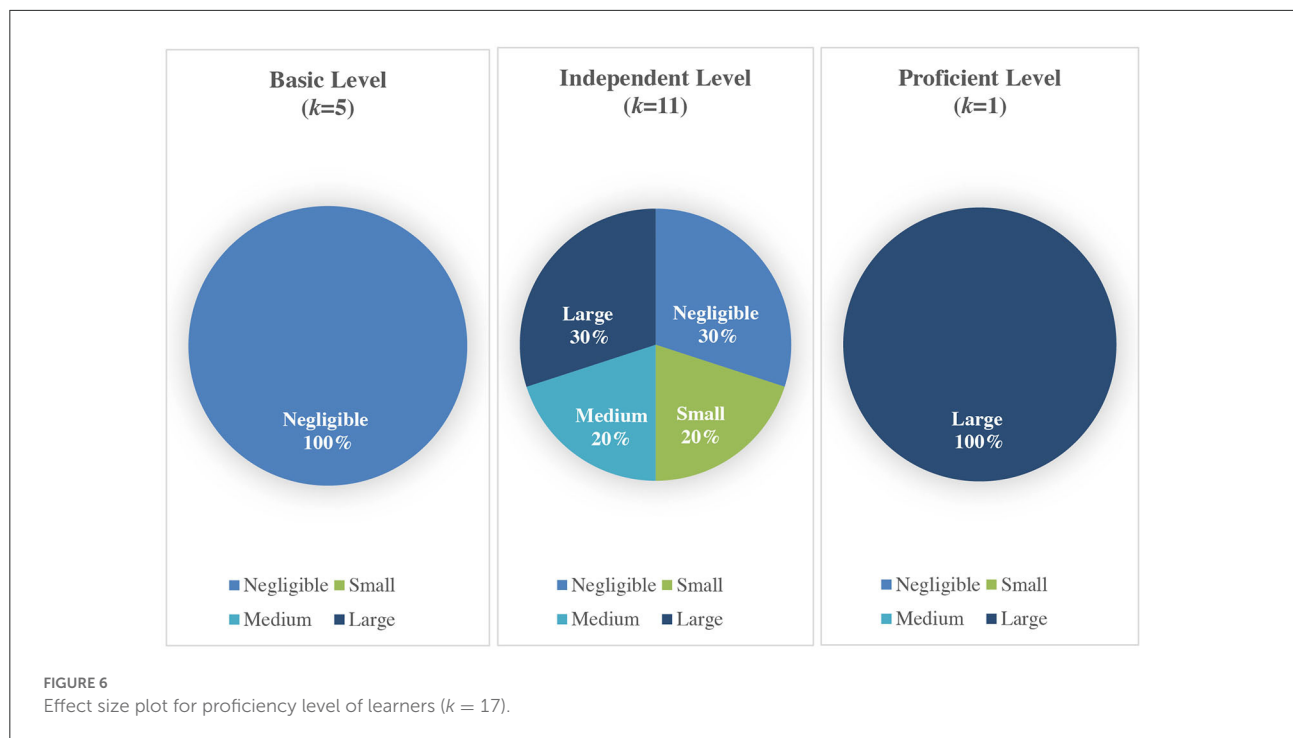
The vocabulary assessments used in the reviewed studies should have assessed the aspects of word knowledge learned by the learners. As the studies included in this synthesis were related to the use of word cards, there should have been a relationship between the assessments in these studies and the use of the word cards. More specifically, Nation (1982, 2013a) suggested a simultaneous presentation of the L1 form of a target

word and its meaning for the first encounter, and then a delayed presentation when using word cards. This is because a retrieval of a target word's form or meaning is necessary for learning to take place. If word cards were used to learn productive knowledge of form, i.e., learners looking at the L1 meaning and trying to recall the L2 written form, it is assumed that the productive knowledge of form would have been assessed by the researchers. Alternatively, if word cards were used to learn productive knowledge of meaning, i.e., learners looking at the L2 written form and trying to recall the L1 meaning, it is assumed that the productive knowledge of meaning would have been assessed by the researchers.

Surprisingly, researchers assessed receptive knowledge more often than productive knowledge in the reviewed studies. This might be due to two reasons. Firstly, the extent that the learners engaged in productive learning when using word cards were limited. Researchers might not have assessed productive knowledge due to learners not often using the word cards for productive learning. If word cards were used properly by the learners, the learning of productive knowledge would occur, as a retrieval of target words' forms or meanings would take place. In many of the reviewed studies, it is less certain if the learners used the word cards in this way, which could be the reason that researchers did not assess productive knowledge. Only some of the studies clearly indicated the learners used the word cards for the learning of productive knowledge (Alhuwaydi, 2020; Lukas et al., 2020; Yüksel et al., 2020).

In contrast, some other studies simply asked the learners to create word cards but gave limited instruction to the learners on how to use those word cards. For example, in Kuo and Ho's (2012) study, learners were presented with the L2 forms of the target words and were required to write the L1 meanings on the other side of the word cards. The learners were then required to share their experience of using the word card strategy in class. However, it might have been possible that the learners did not know how to use the word cards properly to do the retrievals. It should be mentioned that word card creation is not the end of the learning process. Instead, the key to word card use is for learners to look at the L2 word or L1 meaning on one side, and test themselves to see if they can recall the L1 meaning or L2 word on the other side (Komachali and Khodareza, 2012).

There were also studies that engaged learners in picture-word matching activities, but the number of retrievals were limited. For example, in Samad and Makingkung's (2020) study, learners were required to read the target words on the word cards, and then match the pictures with the target words by pasting them in task books. From what was reported in the article, it seemed the learners did not work any further with the target words. Likewise, in Oberg's (2011) study, learners were presented with L2 forms of the target words and were required to draw pictures to represent the target words on the other side of the word cards. Then they required learners to recall the L1 forms of the target words by looking at the



pictures. From what was reported in the article, it appeared that was where the practice with the target words ended. However, previous research has suggested that learners should check their answers rapidly in an easy way by turning over the cards (Özer and Koçoğlu, 2015). The learners should come back to the word cards repeatedly, as this process will provide them with opportunities to encounter the vocabulary (Nation, 2008; Komachali and Khodareza, 2012).

Many of the studies also seemed to indicate learners only used the word cards while they were in class. Learners should also “engage in retrieval activities” outside the classroom in their spare time (Kuo and Ho, 2012, p. 35), as word cards are convenient to carry around.

The second reason for productive knowledge being assessed less often than receptive knowledge is that productive knowledge is more difficult to assess. Although the creation of a productive knowledge test is simple, the marking of the test is potentially difficult. Usually, productive knowledge is assessed through translation (Kuo and Ho, 2012; Fukushima, 2019). However, there may be multiple possible translations for a single target word. Unlike multiple-choice tests, productive knowledge translation tests do not necessarily have a single answer. However, even with multiple possible answers, there are still reliable ways to mark a productive knowledge translation test. Nation (2013b) suggested setting a way of marking any test, i.e., the use of an answer key and a set of rules for dealing with unusual or unexpected answers should be prepared before the marking begins. The key and rules for marking could be used by anyone that is asked to mark such a test. This can provide an easy

way to mark a productive knowledge test. However, if the target words are selected by the learners, which often occurs when the learners are creating their own word cards for the words selected by themselves, it may be difficult for researchers to anticipate the words selected by learners. This would prevent the researchers from being able to create a single key for marking all learners’ assessment outcomes. Therefore, under the circumstances of learners constructing their own word cards, it may not be possible to create a productive knowledge test.

Vocabulary form, meaning, and use did not receive an equal amount of attention from researchers. Most researchers were interested in assessing form and meaning, possibly because the initial learning of vocabulary occurs when a form-meaning connection is made (Nation, 2013a). On the other hand, vocabulary knowledge of use is a more advanced and complex aspect of vocabulary knowledge, and many learners may “lack the opportunities or motivation to use target words” in a short term study (Yang et al., 2021, p. 479). It is less likely that vocabulary knowledge of use would have been mastered by learners and able to be assessed by researchers, that is unless the study is longitudinal (Nation, 2013a).

Researchers used standardized tests or researcher-constructed tests as vocabulary measurements. If learners’ general vocabulary knowledge was to be assessed, standardized tests were used by the researchers. Alternatively, if learners’ knowledge of specific lexical items were to be assessed, researchers constructed special tests for this purpose. It should be mentioned that most standardized tests used in the previous studies measured the breadth of vocabulary knowledge, i.e., “the

vocabulary size of learners” (Azabdaftari and Mozaheb, 2012, p. 48). It is not surprising that the main standardized tests that were used in the reviewed studies, i.e., VST (Nation and Beglar, 2007), UVLT (Webb et al., 2017) and NGSST (Browne et al., 2013), were all designed to assess learners’ receptive knowledge of meaning.

These tests were found to possess the characteristics of reliability, validity and practicality, which are the three major characteristics of a good test (Nation, 2013b). These tests were a reliable measure of vocabulary size due to their adequate sampling of vocabulary items. These tests were valid because they were measuring what they were supposed to measure, i.e., a vocabulary size test should measure learners’ vocabulary size. These tests were practical because they were easy to administer (a computer program can be used), easy to mark (layout of the tests facilitates marking) and easy to interpret (tested words represent the whole population of words from which they were chosen) (Nation, 2013b).

Considering the characteristics of reliability, validity and practicality can also help to explain the researcher-constructed tests used in the reviewed studies. More specifically, receptive knowledge multiple-choice tests are often very practical in terms of marking but could reduce validity, as it should be more valid for a learner to provide an answer than to choose from a range of choices (Nation, 2013b). For example, Oberg (2011) assessed receptive knowledge by having learners take a sentence fill-in-the-blank test with multiple-choice items given. On the other hand, productive tests are often valid but somehow not practical, as this test format might be challenging to mark. For example, Özer and Koçoğlu (2015) assessed productive knowledge by having learners write a composition using target words. Even with the reduction of practicality, it is necessary to assess learners’ productive knowledge gains from the use of word cards. This is because practicality is not as important as reliability and validity in a test (Nation, 2013b). However, this lack of practicality could also be the reason that less productive knowledge tests were used in previous research on word card use.

## Types of word cards

Interestingly, more of the reviewed studies showed a larger effect for the use of paper word cards compared to digital word cards. It should be mentioned that the comparison between paper word cards and digital word cards was an indirect comparison by looking at the ESs of the particular word card type used in the studies. More studies showed a larger effect for the use of paper word cards compared to digital word cards for two reasons. Firstly, the use of paper word cards might be more suitable for young learners (Azabdaftari and Mozaheb, 2012). Paper word cards allow for easier interaction between learners (Komachali and Khodareza, 2012) and do not necessitate the

learning of a computer program to use them (Reynolds et al., 2020). The familiarization with digital word cards requires time and energy, so teachers or technical staff may be needed to assist the learners to use digital word cards (Azabdaftari and Mozaheb, 2012). Secondly, most studies compared the use of paper word cards to other vocabulary learning activities, i.e., gesture-based systems (Wu et al., 2017), wordlists (Kuo and Ho, 2012) or a control group without any intervention (Komachali and Khodareza, 2012), rather than comparing the two types of word cards. As the use of paper word cards and other activities mentioned above were quite different, it is not surprising that paper word cards were found to be more effective than these other activities. Therefore, the results of this synthesis which indicated paper word cards were more effective than digital word cards should be considered with caution.

Certain studies compared the use of digital word cards and paper word cards. They showed no significant difference in the effectiveness of these two types of word cards (Oberg, 2011; Nikoopour and Kazemi, 2014; Chen and Chan, 2019). However, some studies showed digital word cards had a larger effect than paper word cards on vocabulary learning (Başoğlu and Akdemir, 2010; Azabdaftari and Mozaheb, 2012). Unlike the studies that investigated paper word cards in comparison to other activities, these studies compared digital word cards to paper word cards. Under this circumstance, the use of paper word cards and digital word cards were two similar types of activities. Therefore, it is not surprising that there was a non-significant difference between the use of paper word cards and digital word cards. However, there still were some studies that showed a larger effect with the use of digital word cards. For these studies, the additional affordances of digital media such as incorporating sounds (Başoğlu and Akdemir, 2010; Barkat and Aminafshar, 2015; Fukushima, 2019), animations (Barkat and Aminafshar, 2015; Chen and Chan, 2019) and videos (Chen and Chan, 2019) could have been the reason for the better learning outcome, even though previous researchers have suggested that these could be distractions for learners (Chen and Chan, 2019).

More of the reviewed studies showed a larger effect for the use of ready-made cards than self-constructed word cards. Teachers are usually aware of their learners’ proficiencies and could select target words at an appropriate level of difficulty for their learners. It is important for learners to focus on learning vocabulary that is at the right level of difficulty. Learners should focus on learning the most frequent words in a language first (Nation, 2013a). In other words, the first 1,000 words should be learned before the second 1,000 words, and the second 1,000 words should be learned before the third 1,000 words, and so on (Nation, 2013a). In this regard, if a learner had not mastered the first 1,000 most frequent words of English, the learner should not try to learn words from the third 1,000 most frequent words of English.

Teachers are usually aware of the proficiency level of their learners, so they may have been in a better situation to select

the most appropriate target words for learners (Read, 2000). On the other hand, the self-constructed word cards that were created by the learners on their own may have contained target words that were not at the appropriate level, i.e., the words could have been too easy or too difficult for learners. It could have been that the learners in the previous studies were not well-equipped at target word selection. Teachers should give guidance and training on how to select target words that are appropriate for learners, because the most frequent words of English need to be mastered first (Nation, 2013a). It is difficult to know exactly how the learners used the word cards in the previous research, as some studies only explained the steps involved in word card construction (e.g., Chien, 2013) or how the word cards were used (e.g., Kose and Mede, 2018), but not both. Overall, it is difficult to determine the difference in the effectiveness of ready-made word cards and self-constructed word cards, as none of the included studies compared these two types of word cards.

## Use of word cards

### Learning conditions

More of the reviewed studies showed a larger effect for using word cards in an intentional learning condition compared with an incidental learning condition. However, only 2 of the 32 studies used an incidental learning condition (Özer and Koçoğlu, 2015; Reynolds et al., 2020). Therefore, it is premature to conclude that the intentional learning condition would benefit vocabulary learning more than the incidental learning condition when using word cards. It is not surprising that most of the studies used an intentional learning condition as the use of word cards is an intentional learning strategy (Nation, 2013a). Unless the use of the word cards was manipulated by the researchers to create an experiment in incidental learning, it was less likely that use of word card could result in incidental learning.

Although word cards are usually an intentional learning strategy, they can be used for incidental learning purposes. This is important because intentional and incidental learning should complement each other (Nation, 2013a). For example, a learner who reads an article incidentally could come across an unknown word and then record that word on a word card for later review (Reynolds et al., 2020). Except for the first few thousand most common words of English, most vocabulary should be learned incidentally (Lin and Lin, 2019). After a mastery of these most frequent words using intentional learning strategies, the learner can work on increasing their vocabulary size with incidental vocabulary learning strategies (Lin and Lin, 2019).

Interestingly, more of the reviewed studies showed a larger effect for using word cards in a massed learning condition compared with a spaced learning condition. A teacher who trains learners on how to use word cards usually tells the learners to use the word cards in a spaced learning condition (Kuo and Ho, 2012). However, previous research has suggested that when

new words are first introduced to learners, a massed learning condition may be more effective (Uchihara et al., 2019). During this initial learning, learners should work with the word cards in a massed learning condition because that will result in a large number of repeated encounters with the words (Uchihara et al., 2019).

There were still certain studies that found spaced learning led to better but non-significant differences than massed learning (Kuo and Ho, 2012). Thus, spaced learning could still potentially be more effective than massed learning. However, it could be that learners should use massed learning initially and then follow up with spaced learning. Moreover, previous researchers have suggested to increasingly spread out the meetings with newly learned words using a distribution schedule where the repetitions become increasingly further apart (Nation, 2008). Revisiting of previously learned words can strengthen retention of vocabulary knowledge (Nation, 2008).

### Time spent learning from word cards

A positive non-significant weak correlation was found between the time that learners spent using word cards and their posttest vocabulary scores. The positive relationship suggested that the more time learners spent on vocabulary learning from word cards, the more vocabulary they learned. There might be two reasons for the statistically insignificant result. Firstly, there were only 14 studies that provided necessary data that could be extracted for analysis in this synthesis. The relatively small sample size might increase variability, which resulted in a statistically non-significant correlation. Secondly, the lack of a meaningful relationship between the learning time and vocabulary learning may be due to the “limited ability of certain learners to learn effectively” from word cards (Nakata, 2008, p. 3). Even though certain learners spent more time, if they did not use the time efficiently, they might not learn no matter how much time they spent.

Although the correlation was not statistically significant, there is research that has indicated more time spent on using word cards results in more vocabulary learning (Webb et al., 2020). It should be mentioned that what is more important than the overall amount of time that learners spend using the word cards is probably how they use those word cards. For example, learners should be repeatedly coming back to words instead of meeting them all at once, which is often referred to as spaced learning (Nation and Webb, 2011; Nation, 2013a).

Another issue that was unclear in the previous studies was whether the words that the learners worked on were semantically related or not. Previous researchers showed that when learners worked with a new group of words that are semantically related to each other, it could be more difficult to acquire them rather than if these words were not semantically related to each other (Tinkham, 1997; Nation and Webb, 2011). When learners are trying to learn a set of semantically related words, they



will confuse words that are too similar, which could increase the learning difficulty (McDonald and Reynolds, 2021). In contrast, if the words are not related to each other or organized thematically, it may lead to a better learning outcome. This is because differences between lexical items facilitates learning (McDonald and Reynolds, 2021).

## Learner proficiency

Learners at the proficient level (i.e., C1 or C2 level) learned the most vocabulary from using word cards, followed by independent (i.e., B1 or B2 level) learners and finally basic (i.e., A1 or A2 level) learners. In other words, learners that were more proficient in English learned more words from using word cards than those less proficient in English. There is a possibility that learner proficiency has a moderating effect on vocabulary learning. Previous researchers have indicated that vocabulary development progresses differently for learners at different proficiency levels (Elgort, 2017). However, it is a relatively under-researched area in the word card literature, as only one of the included studies involved learners at different proficiency levels in a single study (Tsai, 2018). Tsai (2018) found that the learners at the higher proficiency level had more effective learning outcomes than learners at the lower proficiency level.

Language proficiency could be related to the amount of effort needed to invest in the learning task. Learners with lower proficiencies might have to work very hard to learn, which they might consider as a time-consuming task (Elgort, 2017). However, it might be easier for learners with higher proficiencies to gain more vocabulary knowledge, so they might be more willing to invest more time in learning. More proficient learners have more autonomy to take better charge of their learning (Lin and Lin, 2019), and therefore may be more skillful in vocabulary learning using word cards (de Vos et al., 2018). For example, in a study conducted by Azabdaftari and Mozaheb (2012), proficient learners used the word cards both inside and outside the classroom. This may have allowed them to devote more attention to unknown words and may have increased the potential for vocabulary learning. Learners with lower proficiency might stop using the word cards when class is over, even if they are encouraged to use word cards outside of class.

## Conclusion

### Limitations

Although this synthesis uncovered some interesting findings, some limitations must be discussed. Firstly, a literature search was only conducted in Scopus and WOS (including

SCI-Expanded, SSCI, AHCI, CPCI-S, CPCI-SSH, BKCI-S, BKCI-SSH, ESCI). Thus, the coverage of the synthesis is limited to these databases.

Secondly, this synthesis only included published research. Peer-reviewed studies were selected for review to ensure quality, but this opens up the possibility of publication bias.

Thirdly, a synthesis method was adopted to visually present the effect sizes. The studies included in the synthesis were not implemented with totally identical research designs and did not assess the same aspect of word knowledge. Although ESs were calculated for each included study, we did not look at the significant differences between moderating variables. Instead, the effect direction plots and effect size plots (ESs were categorized by their size) were presented. We took this approach which was different from what would be done with meta-analyses in order to be able to include more primary studies in this synthesis.

Finally, only the immediate posttest data was used for ESs analysis. We only looked at the immediate posttest data because all primary studies provided this data. Therefore, the long-term effects of vocabulary learning from word cards were not investigated in this synthesis. With these limitations in mind, the research implications and teaching implications of the current study results are reported below.

## Research implications

One suggestion for future research is that learner proficiency should be reported. Fifteen of the 32 included studies (46.88%) did not report the proficiency level of the learners. It is difficult to interpret the results of studies that do not clearly describe learner proficiency. Due to the uncertainty of learner proficiency, it is also difficult to conclude whether the vocabulary learning reported in such studies can be generalized to certain learner populations. In addition, learners at different proficiency levels can be recruited for future studies. These studies could compare the effects of word card use on basic, independent, and proficient learners' vocabulary learning.

Another suggestion for future research is that appropriate tests should be adopted to test vocabulary learning performance. For example, future researchers can adopt more standardized tests or report the reliability for researcher-constructed tests. Since most of the included studies that used researcher-constructed tests did not provide any reliability measures for the tests, the effects of the word card intervention reported in the studies is questionable. In addition, productive tests should be used for testing learners' vocabulary knowledge gained from the use of word cards, as productive knowledge production should have taken place during word card use.

Future researchers can also compare the effectiveness of ready-made word cards and self-constructed word cards, as none of the reviewed studies compared the effect of these two

types of word cards. In other words, the effects of both types of word cards had been studied separately but were not compared in a single study.

Another interesting area for future research may be to investigate the learning of different aspects of word knowledge. In this synthesis, most studies started with receptive knowledge in their vocabulary assessments. To prevent practice effects, i.e., previous tests could affect subsequent test performance, so the scores of the first receptive knowledge assessment in each study were extracted for analysis. This synthesis provides a certain understanding of how receptive knowledge of form and meaning can be acquired through the use of word cards. However, we are less certain of the effects of the variables in the current synthesis have on other aspects of word knowledge acquired through the use of word cards. As more time may be needed to develop vocabulary knowledge of use, future longitudinal studies can be conducted to address this gap in the literature.

## Teaching implications

A teacher that decides to incorporate the use of word cards inside or outside their language classroom should take the following into consideration. Whether using digital or paper word cards, teachers should spend adequate time providing guidance to learners on how to use word cards properly (Chen and Chan, 2019). Some digital programs could offer teachers some affordances such as ready-made word banks and streamlined use of the word cards. If a teacher chooses to use digital word cards, large screen tablets or computer should be used, because they can provide a better learning experience. While the synthesis did not aim to investigate how to use digital word cards for better learning outcomes, teachers who plan to use digital word cards should consider the screen size in a computer program or a mobile app. This is because previous researchers (Ji and Aziz, 2021) indicated that learners might have difficulty learning vocabulary on devices with limited screen sizes.

For some learners, self-constructed word cards can save teacher planning time (Reynolds et al., 2020), but require proper guidance before learners begin constructing the word cards (Reynolds and Shih, 2019) and additional checks afterwards by teachers (Reynolds et al., 2020). For self-constructed word cards, teachers should guide the learners in target word selection, especially for those who have just received training on how to use word cards.

A teacher should consider encouraging different learning conditions when learners are using word cards. Intentional learning of vocabulary using word cards is very effective, but teachers can also consider asking learners to use word cards for new words they have incidentally encountered through engagement in other non-language learning tasks, such as reading, watching videos (Lin and Lin, 2019), or classroom discussion (Uchiyama et al., 2019). Teachers should stress the importance of when spaced and massed learning should be

applied to word card use. Teachers can encourage learners to use massed learning when they work with new words at the beginning, and use spaced learning later on. Learners also need to be reminded to use the word cards frequently throughout the day. Many of these guidelines for using word cards have also been incorporated into certain digital apps. However, if a teacher suggests digital apps to learners, the teacher should make sure the digital apps possess these qualities before recommending them to learners.

## Data availability statement

The original contributions presented in the study are included in the article/Supplementary material, further inquiries can be directed to the corresponding author/s.

## Author contributions

BR contributed to the conception and design of the study and revised the paper. YL collected, organized, analyzed the data, and drafted the paper. BR and YL interpreted the results, approved the submitted version of the paper, responded to the reviewer comments, and revised the submitted version of the paper. Both authors contributed to the article and approved the submitted version.

## Funding

This research was supported by the University of Macau under Grant Number MYRG2019-00030-FED.

## Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

## Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

## Supplementary material

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fpsyg.2022.984211/full#supplementary-material>

## References

- \*Alhuwaydi, A. A. (2020). Effect of smartphone flashcard app on Saudi undergraduates' vocabulary acquisition in EFL reading classes. *Asian EFL J.* 27, 227–250.
- \*Aminafshar, N. (2017). EFL language learners' perception for learning vocabulary by using CAP and flash cards. *Modern J. Lang. Teach. Methods* 7, 86–94.
- \*Azabdafatari, B., and Mozaheb, M. A. (2012). Comparing vocabulary learning of EFL learners by using two different strategies: mobile learning vs. flashcards. *Eurocall Rev.* 20, 47–59. doi: 10.4995/eurocall.2012.11377
- \*Barkat, B., and Aminafshar, N. (2015). Effect of call-based and flash card-based techniques of teaching on the target language vocabulary learning. *Modern J. Lang. Teach. Methods* 5, 49–54.
- \*Başoğlu, E. B., and Akdemir, Ö. (2010). A comparison of undergraduate students' English vocabulary learning: using mobile phones and flash cards. *TOJET Turkish Online J. Educ. Technol.* 9, 1–7.
- Boulton, A., and Cobb, T. (2017). Corpus use in language learning: a meta-analysis. *Lang. Learn.* 67, 348–393. doi: 10.1111/lang.12224
- Browne, C., Culligan, B., and Phillips, J. (2013). *New General Service List Project*. Available online at: [www.newgeneralservicelist.org](http://www.newgeneralservicelist.org) (accessed July 24, 2022).
- \*Chen, R. W., and Chan, K. K. (2019). Using augmented reality flashcards to learn vocabulary in early childhood education. *J. Educ. Comput. Res.* 57, 1812–1831. doi: 10.1177/0735633119854028
- \*Chien, C.-W. (2013). "Perception and practice of Taiwanese EFL learners' making vocabulary flashcards on Quizlet," in *IADIS International Conference E-learning 2013*, Eds M. B. Nunes and M. McPherson (Prague: IADIS International Association for Development for the Information Society), 459–462.
- \*Chien, C.-W. (2015). Analysis the effectiveness of three online vocabulary flashcard websites on L2 learners' level of lexical knowledge. *English Lang. Teach.* 8, 111–121. doi: 10.5539/elt.v8n5p111
- Cooper, H. (2017). *Research Synthesis and Meta-Analysis: A Step-by-Step Approach*. Los Angeles: Sage.
- Council of Europe (2001). *Common European Framework of Reference for Languages: Learning, Teaching, Assessment*.
- Crow, J. T. (1986). Receptive vocabulary acquisition for reading comprehension. *Modern Lang. J.* 70, 242–250. doi: 10.1111/j.1540-4781.1986.tb05271.x
- de Vos, J. F., Schriefers, H., Nivard, M. G., and Lemhöfer, K. (2018). A meta-analysis and meta-regression of incidental second language word learning from spoken input. *Lang. Learn.* 68, 906–941. doi: 10.1111/lang.12296
- Elgort, I. (2017). Technology-mediated second language vocabulary development: a review of trends in research methodology. *CALICO J.* 35, 1–29. doi: 10.1558/cj.34554
- Fleckenstein, J., Keller, S., Krüger, M., Tannenbaum, R. J., and Köller, O. (2020). Linking TOEFL iBT® writing rubrics to CEFR levels: cut scores and validity evidence from a standard setting study. *Assess. Writ.* 43, 1–15. doi: 10.1016/j.asw.2019.100420
- \*Fukushima, S. (2019). EmoTan: enhanced flashcards for second language vocabulary learning with emotional binaural narration. *Res. Pract. Technol. Enhanc. Learn.* 14, 1–19. doi: 10.1186/s41039-019-0109-0
- \*Galedari, Z. R., and Basirou, S. R. (2016). The effect of prompt cards on vocabulary learning: comprehension and production. *Modern J. Lang. Teach. Methods* 6, 231–238.
- \*Hamzehbagi, R., and Bonyadi, A. (2015). The effects of flashcards on the EFL high school female students' vocabulary knowledge and reading ability. *Modern J. Lang. Teach. Methods* 5, 232–239.
- \*Hidayat, M. T., and Yulianti, A. K. (2020). The effectiveness of flashcard augmented reality media and game chick learn on the ability to memorize vocabulary in English primary school students. *Int. J. Innov. Creativ. Change* 11, 151–168.
- Ji, P. W., and Aziz, A. A. (2021). A systematic review of vocabulary learning with mobile-assisted learning platforms. *Int. J. Acad. Res. Bus. Soc. Sci.* 11, 1503–1521. doi: 10.6007/IJARBS/v11-i11/11383
- Kim, S. K., and Webb, S. (2022). The effects of spaced practice on second language learning: a meta-analysis. *Lang. Learn.* 72, 1–51. doi: 10.1111/lang.12479
- \*Komachali, M. E., and Khodareza, M. (2012). The effect of using vocabulary flash card on Iranian pre-university students' vocabulary knowledge. *Int. Educ. Stud.* 5, 134–147. doi: 10.5539/ies.v5n3p134
- \*Kose, T., and Mede, E. (2018). Investigating the use of a mobile flashcard application rememba on the vocabulary development and motivation of EFL learners. *Mextesol J.* 42, 1–26.
- \*Kuo, Y., and Ho, H.-Y. (2012). Effects of word card strategy versus word list strategy on Taiwanese EFL junior high school students' vocabulary retention. *Electr. J. Foreign Lang. Teach.* 9, 26–45.
- \*Lai, C.-H., Jong, B.-S., Hsia, Y.-T., and Lin, T.-W. (2020). Integrating flash cards with narratives for mobile learning of English vocabulary. *Int. J. Interact. Mobile Technol.* 14, 4–16. doi: 10.3991/ijim.v14i04.11723
- Larson-Hall, J. (2010). *A Guide to Doing Statistics in Second Language Research Using SPSS*. New York, NY: Routledge.
- Lauffer, B., and Hulstijn, J. (2001). Incidental vocabulary acquisition in a second language: the construct of task-induced involvement. *Appl. Ling.* 22, 1–26. doi: 10.1093/applin/22.1.1
- \*Lavoie, C. (2016). The effect of training on vocabulary strategy use: explicit teaching of word family, word network and word card strategies. *J. Lang. Teach. Learn.* 6, 20–34.
- Light, R. J., and Pillemer, D. B. (1984). *Summing Up: The Science of Reviewing Research*. Cambridge: Harvard University Press.
- Lin, J.-J., and Lin, H. (2019). Mobile-assisted ESL/EFL vocabulary learning: a systematic review and meta-analysis. *Comput. Assist. Lang. Learn.* 32, 878–919. doi: 10.1080/09588221.2018.1541359
- \*Lukas, B. A., Patrick, F. I. A., Chong, G., Jaino, N. B., and Yunus, M. M. (2020). Using U-NO-ME card game to enhance primary one pupils' vocabulary. *Int. J. Learn. Teach. Educ. Res.* 19, 304–317. doi: 10.26803/ijlter.19.5.19
- McDonald, J. A., and Reynolds, B. L. (2021). Learning semantic and thematic vocabulary clusters through embedded instruction: effects on very young English learners' vocabulary acquisition and retention. *Appl. Ling. Rev.* doi: 10.1515/applirev-2020-0102
- Mondria, J. A., and Mondria-De Vries, S. (1994). Efficiently memorizing words with the help of word cards and "hand computer": theory and applications. *System* 22, 47–57. doi: 10.1016/0346-251X(94)90039-6
- Morris, S. B., and DeShon, R. P. (2002). Combining effect size estimates in meta-analysis with repeated measures and independent-groups designs. *Psychol. Methods* 7, 105–125. doi: 10.1037/1082-989X.7.1.105
- \*Nakata, T. (2008). English vocabulary learning with word lists, word cards and computers: implications from cognitive psychology research for optimal spaced learning. *ReCALL* 20, 3–20. doi: 10.1017/S0958344008000219
- Nakata, T. (2011). Computer-assisted second language vocabulary learning in a paired-associate paradigm: a critical investigation of flashcard software. *Comput. Assist. Lang. Learn.* 24, 17–38. doi: 10.1080/09588221.2010.520675
- Nation, I. S. P. (1982). Beginning to learn foreign vocabulary: a review of the research. *RELC J.* 13, 14–36. doi: 10.1177/003368828201300102
- Nation, I. S. P. (2008). *Teaching Vocabulary: Strategies and Techniques*. Boston, MA: Heinle Cengage Learning.
- Nation, I. S. P. (2013a). *Learning Vocabulary in Another Language*. Cambridge: Cambridge University Press.
- Nation, I. S. P. (2013b). *What Should Every EFL Teacher Know?* South Korea: Compass Publishing.
- Nation, I. S. P., and Beglar, D. (2007). A vocabulary size test. *Lang. Teach.* 31, 9–12.
- Nation, I. S. P., and Webb, S. (2011). "Learning vocabulary from word cards," in *Researching and Analyzing Vocabulary* (Heinle, Cengage Learning), 29–44.
- Navarro, D. (2013). *Learning Statistics With R: A Tutorial for Psychology Students and Other Beginners*. Sydney: University of New South Wales.
- \*Nikoopour, J., and Kazemi, A. (2014). Vocabulary learning through digitized and non-digitized flashcards delivery. *Procedia Soc. Behav. Sci.* 98, 1366–1373. doi: 10.1016/j.sbspro.2014.03.554
- \*Ober, A. (2011). Comparison of the effectiveness of a call-based approach and a card-based approach to vocabulary acquisition and retention. *CALICO J.* 29, 118–144. doi: 10.11139/cj.29.1.118-144
- \*Özer, Y. E., and Koçoglu, Z. (2015). "The use of quizlet flashcard software and its effects on vocabulary recall," in *Proceedings of INTED2015 Conference*, Ed M. Kesselman (Madrid: Ankara Üniversitesi Tömer), 1630–1634.

- Page, M. J., McKenzie, J. E., Bossuyt, P. M., Boutron, I., Hoffmann, T. C., Mulrow, C. D., et al. (2021). The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. *Br. Med. J.* 372, 1–9. doi: 10.31222/osf.io/v7gm2
- Paivio, A. (1979). *Imagery and Verbal Processes*. Hillsdale, NJ: Psychology Press.
- Pimsleur, P. (1967). A memory schedule. *Modern Lang. J.* 51, 73–75. doi: 10.1111/j.1540-4781.1967.tb06700.x
- Prion, S., and Haerling, K. A. (2014). Making sense of methods and measurement: Spearman-rho ranked-order correlation coefficient. *Clin. Simul. Nurs.* 10, 535–536. doi: 10.1016/j.ecns.2014.07.005
- Read, J. (2000). “The design of discrete vocabulary tests,” in *Assessing Vocabulary* (Cambridge University Press), 150–187.
- \*Reynolds, B. L., and Shih, Y.-C. (2019). The learning effects of student-constructed word cards as homework for the adolescent English Language classroom. *System* 81, 146–162. doi: 10.1016/j.system.2019.01.005
- \*Reynolds, B. L., Wu, W.-H., and Shih, Y.-C. (2020). Which elements matter? Constructing word cards for vocabulary growth. *SAGE Open* 10, 1–12. doi: 10.1177/2158244020919512
- Sadoski, M. (2006). A dual coding view of vocabulary learning. *Read. Writ. Q.* 21, 221–238. doi: 10.1080/10573560590949359
- \*Samad, F., and Makingkung, V. (2020). “The use of word card media to improve early reading skill at preschool,” in *Proceedings of the International Conference on Industrial Engineering and Operations Management*, Ed Industrial Engineering and Operations Management Society (Dubai: IEOM Society International), 2135–2140.
- \*Saputri, T. (2017). “Improving vocabulary mastery through flashcards in Sartika kindergarten Surabaya,” in *International Conference on English Language Teaching (ICONELT 2017)*, Ed Rakhmawati (Surabaya: Atlantis Press), 214–218.
- Stoeckel, T., Bennett, P., and Ishii, T. (2018). A Japanese-English bilingual version of the New General Service List Test. *JALT J.* 40, 5–21. doi: 10.37546/JALTJJ40.1-1
- \*Tan, A., and Nicholson, T. (1997). Flashcards revisited: training poor readers to read words faster improves their comprehension of text. *J. Educ. Psychol.* 89, 276–288. doi: 10.1037/0022-0663.89.2.276
- Thomson, H. J., and Thomas, S. (2013). The effect direction plot: visual display of non-standardised effects across multiple outcome domains. *Res. Synth. Methods* 4, 95–101. doi: 10.1002/jrsm.1060
- Tinkham, T. (1997). The effects of semantic and thematic clustering on the learning of second language vocabulary. *Second Lang. Res.* 13, 138–163. doi: 10.1191/026765897672376469
- \*Tsai, C.-C. (2018). A comparison of EFL elementary school learners’ vocabulary efficiency by using flashcards and augmented reality in Taiwan. *New Educ. Rev.* 51, 53–65. doi: 10.15804/ner.2018.51.1.04
- Uchihara, T., Webb, S., and Yanagisawa, A. (2019). The effects of repetition on incidental vocabulary learning: a meta-analysis of correlational studies. *Lang. Learn.* 69, 559–599. doi: 10.1111/lang.12343
- Webb, S., Sasao, Y., and Ballance, O. (2017). The updated vocabulary levels test developing and validating two new forms of the VLT. *ITL Int. J. Appl. Linguist.* 168, 33–69. doi: 10.1075/itl.168.1.02web
- Webb, S., Yanagisawa, A., and Uchihara, T. (2020). How effective are intentional vocabulary-learning activities? A meta-analysis. *Modern Lang. J.* 104, 715–738. doi: 10.1111/modl.12671
- Wong, K. M., and Samudra, P. G. (2019). L2 vocabulary learning from educational media: extending dual-coding theory to dual-language learners. *Comput. Assist. Lang. Learn.* 34, 1182–1204. doi: 10.1080/09588221.2019.1666150
- Wright, T. S., and Cervetti, G. N. (2017). A systematic review of the research on vocabulary instruction that impacts text comprehension. *Read. Res. Q.* 52, 203–226. doi: 10.1002/rq.163
- \*Wu, G.-Y., Cheng, I. L., Chew, S. W., Zhu, C.-Y., Hsu, C.-N., and Chen, N.-S. (2017). “English vocabulary learning performance and brainwave differences: the comparison between gesture-based and conventional word-card,” in *Innovations in Smart Learning*, Eds E. Popescu, Kinshuk, M. K. Khribi, R. Huang, M. Jemni, N.-S. Chen, and D. G. Sampson (Singapore: Springer), 199–208.
- \*Wulandari, A., and Musfiroh, T. (2020). Effectiveness of using bilingual and monolingual card in increasing vocabulary for young learners. *Int. J. Sci. Technol. Res.* 9, 2211–2214.
- \*Xodabande, I., Pourhassan, A., and Valizadeh, M. (2021). Self-directed learning of core vocabulary in English by EFL learners: comparing the outcomes from paper and mobile application flashcards. *J. Comput. Educ.* 8, 93–111. doi: 10.1007/s40692-021-00197-6
- Yanagisawa, A., and Webb, S. (2021). To what extent does the Involvement Load Hypothesis predict incidental L2 vocabulary learning? A meta-analysis. *Lang. Learn.* 71, 487–536. doi: 10.1111/lang.12444
- Yang, X., Kuo, L. J., Eslami, Z. R., and Moody, S. M. (2021). Theoretical trends of research on technology and L2 vocabulary learning: a systematic review. *J. Comput. Educ.* 8, 465–483. doi: 10.1007/s40692-021-00187-8
- \*Yüksel, H. G., Mercanoglu, H. G., and Yilmaz, M. B. (2020). Digital flashcards vs. wordlists for learning technical vocabulary. *Comput. Assist. Lang. Learn.* 1–17. doi: 10.1080/09588221.2020.1854312

\* Studies included in this synthesis are marked with an asterisk.



# Frontiers in Psychology

Paving the way for a greater understanding of human behavior

The most cited journal in its field, exploring psychological sciences - from clinical research to cognitive science, from imaging studies to human factors, and from animal cognition to social psychology.

## Discover the latest Research Topics

[See more →](#)

### Frontiers

Avenue du Tribunal-Fédéral 34  
1005 Lausanne, Switzerland  
[frontiersin.org](https://frontiersin.org)

### Contact us

+41 (0)21 510 17 00  
[frontiersin.org/about/contact](https://frontiersin.org/about/contact)

