


ORIGINAL ARTICLE

Incidental learning of English derivational noun suffixes through reading: an eye-tracking study

Jiaxuan Shi¹, Barry Lee Reynolds^{2,3}  and Brian V. Rusk²

¹School of English, University of Hong Kong, Hong Kong SAR, China, ²Faculty of Education, University of Macau, Macau SAR, China and ³Centre for Cognitive and Brain Sciences, University of Macau, Macau SAR, China

Corresponding author: Barry Lee Reynolds; Email: barryreynolds@um.edu.mo

(Received 21 January 2024; revised 24 May 2025; accepted 20 June 2025)

Abstract

Previous research has shown that second language (L2) learners can learn new words incidentally through contextual clues. However, little is known about whether derivational affixes can be learned in a similar manner. Addressing this gap, the current study examined whether English as a Foreign Language (EFL) learners could acquire knowledge of English derivational noun suffixes through contextualized reading. Forty Chinese EFL learners participated in the study, completing offline pre-tests and post-tests to assess improvements in three aspects of suffix knowledge. Participants' eye movements during reading were also recorded to investigate the relationships between online processing of derived words and suffix learning. The offline test results showed that the learners made significant progress in three aspects of suffix knowledge. Multilevel logistic regression analyses further indicated that improvements in accuracy were significantly predicted by eye fixation measures, learners' L2 proficiency, and language-level factors. Findings indicate that incidental learning while reading can effectively supplement intentional learning, particularly for English affixes that occur less frequently.

Keywords: Affix knowledge; Chinese EFL learners; English derivational suffixes; eye tracking; incidental learning

Introduction

Vocabulary knowledge is vital to the successful acquisition of a second language (L2). Typically, L2 learners begin by intentionally learning vocabulary from dictionaries, textbooks, or teacher instruction (Hulstijn, 2001). Once they have acquired the most commonly used words, they transition to becoming independent learners, relying primarily on incidental vocabulary acquisition through reading (Nagy et al., 1987). Many studies have examined how reading in context promotes the incidental learning

© The Author(s), 2025. Published by Cambridge University Press. This is an Open Access article, distributed under the terms of the Creative Commons Attribution-NonCommercial-NoDerivatives licence (<https://creativecommons.org/licenses/by-nc-nd/4.0/>), which permits non-commercial re-use, distribution, and reproduction in any medium, provided that no alterations are made and the original article is properly cited. The written permission of Cambridge University Press must be obtained prior to any commercial use and/or adaptation of the article.

of L2 vocabulary (e.g., Webb, 2008). However, this body of research has largely overlooked the acquisition of affix knowledge—a critical component of word knowledge. Affix knowledge has been found to correlate with L2 learners' overall vocabulary knowledge (Schmitt & Meara, 1997; Mochizuki & Aizawa, 2000; Sasao & Webb, 2017; Laufer *et al.*, 2021; Snoder & Laufer, 2022) and their ability to produce derivatives (Iwaizumi & Webb, 2023). Like individual words, there are numerous affixes that L2 learners need to acquire to achieve language mastery (Bauer & Nation, 1993). As with individual lexemes, affixes vary in frequency; even those that occur less frequently are necessary for English proficiency (Bauer & Nation, 1993). Similar to lexical acquisition, it should be expected that L2 learners initially focus on mastering the most frequently used affixes through intentional learning, then transition to incidental means for the low-frequency ones. It has been suggested that high-frequency affixes are often explicitly taught in L2 classrooms (Nation, 2022). However, the extent to which L2 learners can acquire knowledge of lower-frequency affixes incidentally through reading remains unclear.

Eye-tracking studies investigating incidental vocabulary learning have examined how L2 learners process novel words during reading (Godfroid *et al.*, 2013; Pellicer-Sánchez, 2016; Mohamed, 2018; Elgort *et al.*, 2018; Pellicer-Sánchez, 2021). These studies consistently indicate that learning gains are positively predicted by the duration of eye fixations on novel words, suggesting that greater attention facilitates more substantial learning. However, little is known about whether this association between reading time and learning generalizes to affix knowledge. Given this research gap, the current study investigates whether L2 learners can incidentally learn affix knowledge through reading. Specifically, the study assesses gains in three types of affix knowledge—form recognition, meaning recall, and meaning recognition—using pre-tests and post-tests conducted before and after reading. Eye tracking, which has been applied to explore the effects of morphemic unit properties (e.g., suffix length and the informativeness difference between base words and suffixes) on the processing and reading of derived words (Kuperman *et al.*, 2010), was employed to examine the relationship between the processing of affixed words during reading and incidental learning. Furthermore, affix characteristics and learners' L2 proficiency were analyzed to determine their contributions to learning gains. If the results show that L2 learners can incidentally learn novel affixes through reading, this would have significant implications for both L2 learning and teaching practices.

Background literature

Incidental L2 vocabulary learning through reading

Incidental learning is commonly intended to refer to the process by which language learners acquire knowledge of a language unintentionally through meaning-focused activities like reading or listening, whereas intentional learning involves a deliberate and effortful focus on language form (e.g., Hulstijn, 2001). However, in practice, it is often difficult to determine the moment-to-moment nature of learning (Webb, 2020), and therefore, it is often operationalized as either (a) learning from engagement with meaning-focused tasks (e.g., Ellis, 1999; Chen & Truscott, 2010) or (b) learning prior to an unannounced post-test (e.g., Hulstijn, 2001). In the field of

L2 acquisition, many studies have explored incidental vocabulary learning (e.g., Webb, 2008; Elgort & Warren, 2014; Pellicer-Sánchez, 2016). Though some researchers have reported that the vocabulary gains from incidental learning during reading are modest (Hulstijn et al., 1996), this approach is still widely regarded as an important complement to intentional vocabulary learning (Schmitt & Carter, 2000; Wu, 2009; Hu, 2013). Drawing on Nation's (2013) classification of word knowledge, studies on L2 incidental vocabulary learning typically assess learners' gains using tests of word form recognition, meaning recall, and meaning recognition (e.g., Pellicer-Sánchez, 2016).

Although the potential benefits of incidental learning through reading are influenced by various factors, including individual learner differences, learning materials, activity types, and methodological control for prior knowledge (Webb et al., 2023), the number of repeated exposures to a new word remains the most commonly researched factor (Horst et al., 1998; Waring & Takaki, 2003; Pellicer-Sánchez & Schmitt, 2010; Pellicer-Sánchez, 2016). Studies have found that the chances of learning a new word increase with repeated exposures, with the number required for incidental word learning ranging from as few as two to more than 10. However, the benefit of additional exposures appears to decline beyond 20 exposures (Uchihara et al., 2019). While there is no consensus on the exact number, research by Horst et al. (1998), Waring and Takaki (2003), and Pellicer-Sánchez (2016) suggests that a minimum of eight exposures is necessary for learners to recognize or recall the meaning of new words.

Reading context is also crucial in investigating the incidental learning of words through reading (e.g., Webb, 2008). Previous research has examined various aspects of context, including differences in length, genre, authenticity, and semantic diversity (Pellicer-Sánchez & Schmitt, 2010; Godfroid et al., 2013; Elgort & Warren, 2014; Pellicer-Sánchez, 2016; Johns et al., 2016; Joseph & Nation, 2018). A core concept related to reading context, *contextual richness*, refers to the extent of information a context provides about the meaning of an unknown word (Zahar et al., 2001; Webb, 2008). Some scholars (Zahar et al., 2001; Webb, 2008) have developed a 4-point scale to measure the richness of a word's context. However, research findings on the effect of contextual richness on incidental word learning are inconsistent. While some studies suggest that contextual richness is more influential than exposure frequency (Webb, 2008), others have found the opposite (Zahar et al., 2001; Tekmen & Daloglu, 2006; Reynolds, 2020).

In sum, incidental learning during reading can contribute to multiple aspects of vocabulary knowledge, which are influenced by factors like exposure frequency and contextual richness. Thus, studies on incidental vocabulary learning should account for these factors and employ diverse tests to assess different dimensions of vocabulary knowledge.

L2 Learning and processing of derivational affixes

English words can be modified by a range of inflectional and derivational morphemes, with derivational affixes often changing a word's meaning and/or part of speech. For instance, adding the derivational prefix *un-* and suffix *-ness* to *happy* creates *unhappiness*, altering both its meaning and grammatical category.

Previous research has explored various factors that may influence the learning and processing of L2 affixes. One aspect is the distinction between neutral and non-neutral affixes regarding phonological changes. Non-neutral affixes alter either stress placement or vowel quality in the base word, such as the suffix *-ese* in *Japanese*. In contrast, neutral affixes, such as *-hood* in *childhood*, leave the base word's stress and vowel quality unchanged. This distinction potentially influences acquisition, as neutral affixes are generally learned earlier and more easily by L1 learners (Tyler & Nagy, 1989) and likely by L2 learners as well (Friedline, 2011). Another aspect is affix productivity, which also appears to influence the processing and acquisition of derived words (e.g., Bertram *et al.*, 1999; Bertram *et al.*, 2000). For example, native Finnish children better recognize and understand words with highly productive suffixes (Bertram *et al.*, 2000). Similarly, affix frequency is considered critical for developing morphological knowledge in children (e.g., Beyersmann *et al.*, 2012; Rastle & Davis, 2008), and has been examined for its role in nonword priming among native French speakers (Beyersmann *et al.*, 2015). Furthermore, affix position—prefix or suffix—may affect L1 readers' processing, potentially stemming from left-to-right parsing during reading (Beauvillain, 1996; Kwanten & Mewhort, 1999; Mousikou & Schroede, 2019). The part of speech of derived words also presumably impacts learning. Schmitt and Zimmerman (2002) found that nouns or verbs were easier to produce compared to adjectives or adverbs, although Iwaizumi and Webb (2023) found that producing adverb derivatives was the easiest.

Learners' language proficiency also influences L2 affix learning and processing. Research has shown a positive correlation between L2 proficiency and affix knowledge (Mochizuki & Aizawa, 2000; Laufer *et al.*, 2021; Snoder & Laufer, 2022; Laufer, 2023), with advanced learners more likely to possess richer knowledge of derivational morphemes. Their broader vocabulary and derivational affix knowledge may enable them to acquire derived words more systematically (Iwaizumi & Webb, 2023). This indicates that advanced L2 learners tend to use decompositional processing for morphologically derived words, analyzing words into their constituent morphemes (Diependaele *et al.*, 2011; Li *et al.*, 2017). Furthermore, learners' morphological awareness may also impact L2 affix learning by enhancing their ability to recognize, reflect on, and manipulate the morphemic structure of words while reading (Carlisle, 1995). Several morphological awareness measures have been employed to assess learners' knowledge of derivational morphology (e.g., Mahony, 1994; Mahony *et al.*, 2000; Friedline, 2011). According to Tyler and Nagy (1989), knowledge of derivational morphology encompasses syntactic, relational, and distributional components. Syntactic knowledge helps learners recognize that derivational suffixes represent the syntactic category of words (Tyler & Nagy, 1989, p. 649). Relational knowledge involves understanding that words sharing the same base have related lexical meanings, while distributional knowledge pertains to comprehending the selectional constraints on affixes (Tyler & Nagy, 1989).

To investigate the ongoing debate about L2 morphological processing (i.e., *whole-word retrieval* vs. *morphological decomposition*), researchers often use pseudowords to explore how L2 learners process derivational morphology (e.g., Casalis *et al.*, 2015; Li *et al.*, 2017; Amenta *et al.*, 2025). Pseudowords are novel linguistic forms created by manipulating constituent morphemes (i.e., base words and affixes) for experimental purposes (e.g., *proudy*, *foodle*; see further

examples below). These pronounceable, word-like constructs conform to the phonological and orthographic rules of a given language but lack semantic meaning (Ziegler et al., 1997). Novel-derived words (or *affixed pseudowords*) are a specific subtype formed by combining a real base word with a real affix (e.g., *animalful*; Amenta et al., 2025), and are therefore interpretable to some extent. Although recent studies have shown that both novel-derived words (e.g., *animalful*) and uninterpretable pseudowords (e.g., *animalfil*) may elicit semantic activation (Marelli & Baroni, 2015; Gatti et al., 2023; de Varda et al., 2024), pseudowords remain a valuable tool for examining L2 morphological processing, as any morphological effects observed with pseudowords likely reflect processes that are engaged when whole-word access is unavailable (Li et al., 2017).

For example, Casalis et al. (2015) created four types of pseudowords by orthogonally manipulating the base word and suffix status: (1) a real base word with a real suffix (e.g., *proudy*), (2) a real base word with a non-existent suffix (e.g., *foodle*), (3) a non-existent base word with a real suffix (e.g., *slinny*), and (4) a non-existent base word with a non-existent suffix (e.g., *birtle*). French-speaking L2 learners of English performed a lexical decision task, where affixed pseudowords (e.g., *proudy*) were the hardest to reject, followed by pseudowords with either a real base word or suffix, with fully non-existent forms being the easiest to dismiss. Similarly, Li et al. (2017) tested Chinese-English bilinguals using affixed pseudowords (e.g., *animalful*) and their non-affixed counterparts (e.g., *animalfil*). Participants found affixed pseudowords more difficult to reject, indicating increased morphological processing. Recently, Amenta et al. (2025) examined native Dutch learners of Italian using a lexical decision task with three types of Italian pseudowords: (1) a real base word with a real suffix (e.g., *stelloso*: “with many stars”), (2) a real base word with a non-existent suffix (e.g., *calzeccia* from *calz*- “sock”), and (3) a non-existent base word with a real suffix (e.g., *tencapabile* with *-abile* “-able”). They also found that affixed pseudowords were the hardest to reject. Collectively, these findings suggest that L2 learners are sensitive to derivational morphology, likely due to their reliance on morphological decomposition in the absence of whole-word lexical entries (Li et al., 2017). This process appears to involve activating individual morphemes in lexical memory and evaluating the plausibility of their combination (Li et al., 2017), supporting the view that L2 learners engage in morpheme-level analysis when processing unknown derivations.

Understanding both the form and meaning of affixes is essential for developing affix knowledge (e.g., Bauer & Nation, 1993). Various factors influence an L2 learner’s processing and acquisition of this knowledge, including the properties of the affix itself and individual learner differences. Therefore, research on affix learning should address both language-level factors, such as affix neutrality, and individual-level factors, including L2 proficiency and morphological awareness. Given the advantages of pseudowords in investigating L2 morphological processing, this study employs pseudowords as target stimuli.

Incidental L2 vocabulary learning and eye tracking

Eye fixation measures are commonly used in L2 incidental vocabulary learning research to examine language processing and learning during reading, and are

primarily interpreted as indicators of attention (e.g., Godfroid *et al.*, 2013; Pellicer-Sánchez, 2016; Elgort *et al.*, 2018; Godfroid *et al.*, 2018; Mohamed, 2018; Pellicer-Sánchez *et al.*, 2021). One of the earliest studies was conducted by Godfroid *et al.* (2013), in which 28 Dutch-speaking EFL learners with similar English proficiency were recruited to read paragraphs embedded with 12 targeted pseudowords. The researchers found that learners fixated longer on unknown pseudowords than on known controls, even though both were matched for length and syllable count. Furthermore, the study revealed a positive correlation between the total reading time participants spent fixating on pseudowords and their scores on an unexpected offline recognition test.

Similarly, Pellicer-Sánchez (2016) used eye fixation data to compare the incidental vocabulary learning processes of L1 and L2 English speakers. Twenty-three advanced L2 learners and 25 natives read a story embedded with six targeted pseudowords and six known controls, each appearing eight times. The pseudowords and control words were matched for the number of letters and syllables. Eye movement measures showed that L1 readers spent less time processing the pseudowords than L2 readers. Furthermore, both groups showed a decrease in fixation time for pseudowords and control words with increased exposure. After eight exposures, fixation times did not significantly differ between pseudowords and control words for either group, supporting previous findings that eight exposures to a novel word may be necessary for initial learning to occur (e.g., Horst *et al.*, 1998; Waring & Takaki, 2003). Unannounced offline meaning recall tests revealed a significant positive relationship between total reading time and test scores for both groups.

Mohamed (2018) investigated the processing of 20 unknown pseudowords and 20 control words by 42 advanced L2 learners of English while reading a graded-reader novel. All pseudowords were matched with control counterparts for length and syllable count. Similar to previous studies, Mohamed found that as the occurrences of both pseudowords and control words increased, total reading time spent on these words decreased. Moreover, the results indicated that total reading time positively predicted all three types of vocabulary knowledge. However, exposure frequency had varying impacts on each type of knowledge: form recognition was most affected, followed by meaning recognition, and finally, meaning recall.

In the three aforementioned studies, total reading time on target (pseudo)words strongly and positively predicted increased accuracy on offline vocabulary tests. Each study accounted for L2 proficiency and word-level factors, such as word length. However, none of the studies reported on contextual richness. Previous reading studies have shown that an informative context increases the likelihood of participants inferring the meaning of a target word (Webb, 2008). If so, contextual richness may influence eye fixation patterns: when a word's meaning is inferred before fixation, the time spent fixating on the word may decrease, or the word may even be skipped entirely. Therefore, it is crucial to include ratings of contextual richness in studies using eye-tracking data.

The present study

Previous research has shown that English affix knowledge significantly contributes to the success of L2 English learning. However, limited studies have investigated

whether L2 learners can establish knowledge of English affixes incidentally through reading. To address this gap, the present study employed eye tracking to examine the correlation between the processing of affixed words while reading and incidental learning outcomes. Given the numerous factors that potentially affect affix processing and learning, this study focused on a specific type of English affix: neutral derivational noun suffixes. Learning outcomes were assessed using offline pre-tests and post-tests measuring three types of affix knowledge: suffix form recognition, suffix meaning recognition, and suffix meaning recall. By integrating online eye movement data with offline assessment, the study aimed to provide a clearer understanding of how L2 learners acquire English derivational noun suffixes incidentally while reading. It also analyzed the effects of stimulus-related variables (suffix length, suffix productivity, suffix frequency, and contextual richness) and learner-related variables (vocabulary knowledge and morphological awareness for derivations) on suffix processing and learning. The following research questions were investigated:

1. What aspects of suffix knowledge (i.e., form or meaning) can be incidentally learned by Chinese EFL learners through reading?
2. To what extent do attentional eye fixation measures (i.e., total reading time), language-level characteristics (i.e., suffix length, suffix productivity, suffix frequency, and contextual richness), and learners' language proficiency (i.e., vocabulary knowledge and morphological awareness for derivations) predict learning gains in suffix knowledge through reading?

Given the potential for L2 vocabulary knowledge to be acquired incidentally through reading (Webb, 2008; Pellicer-Sánchez, 2016; Webb et al., 2023), we formulated the following hypotheses:

For Research Question 1, we hypothesized that both suffix form and suffix meaning would be incidentally learned by L2 learners, as reflected in increased accuracy on post-tests compared to pre-tests, across measures of form recognition, meaning recall, and meaning recognition.

For Research Question 2, we hypothesized that longer reading times, indicative of greater attention to targeted base words and suffixes, would predict learning gains. This prediction aligns with prior findings on the positive relationship between increased reading times and incidental vocabulary learning (e.g., Godfroid et al., 2013). Additionally, we expected that stimulus-related factors, including suffix length, suffix productivity, suffix frequency, and contextual richness, would influence the processing and learning of morphologically derived words. Specifically, we anticipated that more informative contexts would facilitate learning, while suffix characteristics like length, productivity, and frequency would also play significant roles. Finally, we predicted that learner-related factors, such as higher levels of morphological awareness for derivations, would lead to greater learning gains due to a better understanding of the morphemic structure of complex words (Carlisle, 1995). Given the close relationship between language proficiency and morphological knowledge (Kraut, 2015), we also hypothesized that vocabulary knowledge, as a measure of language proficiency, would positively impact learning gains.

Methodology

Participants

A total of 52 Chinese EFL learners pursuing non-English related master's degrees at a university in Macau were recruited for this study. Of these, 11 participants were excluded primarily due to technical or calibration issues with the eye-tracking apparatus. Moreover, data from one participant were removed due to low accuracy on true/false reading comprehension questions (below 70%). Therefore, data from 40 participants (8 males; age range: 21–29, $M = 23.8$, $SD = 1.522$) were analyzed. Participants' English proficiency, based on IELTS (International English Language Testing System) or CET-6 (College English Test-6) scores, was estimated at the B2 level (high intermediate) according to the Common European Framework of Reference for Languages (CEFR, Council of Europe, 2001). Results from the updated Vocabulary Levels Test (Webb *et al.*, 2017) confirmed that all participants had receptive knowledge of the most frequent 2,000 English word families. Each participant provided written informed consent before participation and received monetary compensation (approximately USD 15). All participants had normal or corrected-to-normal vision.

Materials

In preparation for the main reading study, targeted pseudowords were created by combining pseudo base words with actual English neutral derivational suffixes. Three-sentence reading contexts were then developed to embed each targeted pseudoword. Pilot testing was conducted to ensure the unfamiliarity of the suffixes to participants in the main study and to gauge the contextual richness of the reading materials. Details of this process are outlined in the following subsections.

Targeted suffixes

To identify potential targeted suffixes, we consulted four resources and compiled a pool of 25 potential English derivational noun suffixes. These suffixes were selected based on examples listed in the following references: (1) The Word Part Levels Test (Sasao & Webb, 2017); (2) infrequent English affixes from Bauer and Nation's (1993) Level 5 word family list; (3) *English Word-Stress* (Fudge, 1984); and (4) *English Suffixes, Stress-assignment Properties, Productivity, Selection and Combinatorial Processes* (Trevian, 2015).

To prepare for the formal study, we created form recognition, meaning recall, and meaning recognition tests for the 25 potential noun suffixes, following a format similar to that of the main study. These tests were administered sequentially—in the order of form recognition, meaning recall, and meaning recognition—to 15 pilot participants with backgrounds equivalent to those of the main study participants. Suffixes recognized by the pilot participants were excluded as potential targets. To ensure accurate eye movement data collection, suffixes of only one or two letters were excluded. This decision was made because evidence indicates that larger interest areas are more likely to be fixated and less likely to be skipped in eye-tracking studies (Conklin & Pellicer-Sánchez, 2016). After applying these criteria, eight suffixes remained, four of which were four letters long, and four of which were

Table 1. Twenty-five potential suffixes and the eight targeted suffixes

Type	Suffix
Twenty-five potential suffixes	-er; -ling; -fest; -ster; -hood; -ess; -ery; -ism; -cade; -ware; -ment; -nik; -ist; -mas; -ate; -tide; -ant; -ure; -ness; -cy; -scape; -ite; -age; -trix; -dom
Eight targeted suffixes	-ling; -ster; -cade; -fest; -nik; -ery; -ate; -ure

three letters long. The full set of 25 potential suffixes and the final eight targeted suffixes are listed in Table 1.

Each of the eight selected suffixes represented multiple lexical meanings. For example, the suffix “-ster” can denote (1) “a person linked with an activity” (e.g., *dopester*, *fraudster*, *gamester*, *gangster*, etc.) or (2) “a conveyance on which or by which people or things are transported” (e.g., *dragster*, *speedster*, *sportster*, etc.) (Trevian, 2015, p. 134). Drawing on the aforementioned resources, one designated meaning was selected for each suffix, ensuring no overlap in meaning across the targeted suffixes. This approach was guided by two considerations: first, the study assesses participants’ learning of suffix meaning knowledge, including meaning recall and meaning recognition; second, using a single, designated meaning minimizes participant confusion. Details about the eight targeted suffixes and their associated information are provided in Appendix A.

Affix productivity and frequency are critical factors in affix learning and processing (Bertram et al., 2000; Beyersmann et al., 2015). Using the *Corpus of Contemporary American English* (COCA), we calculated the total token counts for each targeted suffix to determine suffix frequency. Affix productivity was calculated by comparing the number of *hapax legomena*—types that occur once only in the corpus—with the total number of tokens (i.e., frequency) (Baayen, 1994). The productivity and frequency values for each targeted suffix are presented in Table 2.

Pseudo base words

To conduct the experiment, pseudowords composed of pseudo base words and targeted suffixes were used. Since pseudowords lack lexical meaning, this design prevented participants from inferring suffix meanings from whole words. Moreover, using pseudo base words allowed each occurrence of a suffix to pair with a different pseudo base, thereby promoting the learning of suffixes rather than entire pseudowords.

Previous research suggests that eight exposures may be sufficient for L2 learners to incidentally learn novel words (Pellicer-Sánchez, 2016). Therefore, each targeted suffix was paired with eight unique pseudo base words. These pseudo base words were monosyllabic and five letters in length. The choice of five-letter pseudo base words was informed by eye-tracking research, which indicates that words approximately eight letters or longer are more likely to be fixated (Daneman & Carpenter, 1983). As the targeted suffixes in this study are no less than three letters long, the pseudo base words required a minimum of five letters to maintain the desired word length. Furthermore, all pseudo base words were monosyllabic, as some suffixes could only attach to one-syllable bases.

Monosyllabic nonwords limited to five letters were generated using the ARC Nonword Database (Rastle et al., 2002). To control for orthographic frequency

Table 2. Number of hapax legomena, frequency, and productivity of all the targeted suffixes

Suffix	Number of hapax legomena in COCA	Suffix frequency	Suffix productivity
-ling	6	3625	0.002
-ster	0	3302	0
-cade	3	1325	0.002
-fest	61	1119	0.055
-nik	1	215	0.005
-ery	36	47588	0.001
-ate	0	10238	0
-ure	2	19000	0

effects, pseudo base words were selected based on bigram frequency and OLD20 (Orthographic Levenshtein distance) values (e.g., Mousikou & Schroeder, 2019). Bigram frequencies were obtained from Medler and Binder’s *Online Orthographic Database* (Medler & Binder, 2005). The frequency of the bigram straddling the pseudo base word and suffix was also considered, as it influences readers’ ability to decompose morphologically complex words (Hay, 2002, 2003). The OLD20 value (Yarkoni et al., 2008) was calculated as the average edit distance (insertion, deletion, and substitution) of the 20 nearest neighbors in the lexicon, with higher OLD20 values indicating sparser orthographic neighborhoods (Mousikou & Schroeder, 2019). Wuggy (Keuleers & Brysbaert, 2010) was used to obtain OLD20 values for the generated pseudo base words.

Eight pseudo base words were randomly selected for each targeted suffix, referencing bigram frequencies and OLD20 values, resulting in a total of 64 targeted pseudowords. To ensure comparability across the eight suffix groups, three one-way Analyses of Variance were performed to examine potential differences in bigram frequency within the base word, bigram frequency at the boundary between the pseudo base word and suffix, and OLD20 values. Results showed no significant differences in bigram frequency within the base word ($F(7,56) = 0.044, p = 1.000$), bigram frequency at the boundary ($F(7,56) = 0.069, p = 0.999$), or OLD20 values ($F(7,56) = 0.848, p = 0.552$). These findings suggest that the pseudo base words in each suffix group are comparable across these key orthographic measures.

Since the pseudo bases lack semantic meaning, it was necessary to assign meanings consistent with their assigned suffixes. After finalizing the forms of 64 pseudowords, each was given a meaning that reflected its suffix. For instance, the suffix *-ster*, which conveys the meaning “a conveyance/vehicle,” was used in the pseudoword *wroulster* (composed of *wroul* and *-ster*), which was assigned the meaning “tram” to align with the suffix’s meaning.

Reading contexts

Three-sentence contexts were created for each targeted pseudoword. Each context consisted of a 12-word first sentence and an 8-word last sentence framing the second sentence, which contained the targeted pseudoword. The second sentence

was 22 words long, with the pseudoword consistently appearing as the fifth word. A total of eight neutral sentence frames were used to separate the processing of the targeted pseudoword/context from sentence integration and wrap-up during the eye tracking experiment (Lowell & Morris, 2017). These 64 three-sentence contexts formed the reading materials for this study. To ensure readability, all sentences were written using the first 2,000 most frequent English word families from the British National Corpus/Corpus of Contemporary American English word lists (Nation, 2012). This was confirmed using the *Range for texts* function on the Compleat Lexical Tutor website (Cobb, n.d.). Each three-sentence context underwent independent review and revision by two native English speakers with backgrounds in English language education.

Given that the study focused on incidental suffix learning through reading, it was essential to measure the informativeness of the reading contexts. For this purpose, 15 graduate students with similar demographic backgrounds to those of the main study participants rated the contextual richness of the second sentences. The raters used Webb's (2007) three-level scale for contextual richness, excluding the misleading context option, as none of the contexts were intended to mislead participants. Thus, only the following three ratings were used:

1. It is unlikely that the exact meaning of the target word can be inferred. However, information in the context may lead to partial knowledge of the target word's meaning.

2. Information in the context may make it possible to infer the meaning of the target word. However, there are a number of choices. Participants may gain partial knowledge.

3. Participants have a good chance of inferring the meaning correctly. There are few meanings that are logical apart from the correct meaning. Participants should gain at least partial knowledge (Webb, 2007, p. 53).

The raters were provided with the meanings of the targeted suffixes and the second sentences containing the pseudowords in the fifth word position. Targeted suffixes were bolded, and the informative parts of the context were underlined. Raters rated the extent to which the underlined portions supported inference of the suffix meanings. Task details are provided in Appendix E. A mean contextual richness score was calculated for each context by averaging the raters' scores. These scores for each targeted suffix are presented in Table 3.

Furthermore, 48 filler contexts were created following the same structure and vocabulary constraints. Like the experimental materials, these fillers were composed using the 2,000 most frequent word families (Nation, 2012). To ensure participants paid attention during the reading process, 24 true/false comprehension questions were included, following established methods for assessing reader engagement (e.g., Pellicer-Sánchez, 2016; Godfroid et al. 2018).

Offline tests

Suffix tests

To evaluate participants' suffix knowledge acquired through reading, a pre-test and post-test design was employed, with the tests administered one week apart. Both tests consisted of three components: form recognition, meaning recall, and meaning

Table 3. Contextual richness of all the targeted suffixes

Suffix	Contextual richness
-ling	2.200
-ster	2.325
-cade	2.183
-fest	2.383
-nik	2.375
-ery	2.292
-ate	2.383
-ure	2.333

Note: This table shows the mean contextual richness (i.e., informativeness) for each suffix's three-sentence reading context.

recognition, each assessing a different aspect of suffix knowledge. These three suffix tests were divided into two parts. Part one, administered first, assessed form recognition and meaning recall. Form recognition was evaluated using a multiple-choice format similar to the Word Part Levels Test (Sasao & Webb, 2017), where participants were required to select the real English suffix from four options. For meaning recall, a blank space was provided to the right side of the multiple-choice item, where participants were instructed to write the Chinese translation of the selected suffix. In part two, meaning recognition was assessed using a similar multiple-choice format. Each item presented four options matched for part of speech and length, requiring participants to select the correct meaning of the suffix. The test was divided into two parts, as the multiple-choice options in part two could aid participants in answering the meaning recall section of part one. To mitigate the risk of a practice effect, the pre-tests included 17 distractors in addition to the eight targeted suffixes. In the post-test, only knowledge of the eight targeted suffixes was assessed, as no delayed post-test was arranged. Each correct response was assigned one point, with a maximum total score of 24. Pre-test and post-test materials are provided in Appendices B and C, respectively.

Language proficiency tests

To be included, participants were required to demonstrate mastery of the first 2,000 English word families as measured by the Updated Vocabulary Levels Test (Webb et al., 2017). This criterion ensured participants could comprehend the study stimuli. Language proficiency, often measured by vocabulary knowledge, has been suggested to be closely related to morphological knowledge (Kraut, 2015). Therefore, we administered the LexTALE test (Lemhöfer & Broersma, 2012) to assess participants' L2 proficiency.

Participants' knowledge of the three aspects of derivational morphology proposed by Tyler and Nagy (1989), namely syntactic knowledge, relational knowledge, and distributional knowledge, was also evaluated using separate tests. Syntactic knowledge was assessed using an adapted version of the affix-choice

Table 4. Participant information

Variable	<i>M</i>	<i>SD</i>
Age (years)	23.800	1.522
LexTALE	57.219	9.395
Morphological awareness	58.625	5.601

Note: Maximum morphological awareness score is 74. LexTALE score is used as a proxy for language proficiency, with a maximum score of 100.

nonword task developed by Mahony (1994). This cloze-like task required participants to select one out of four suffixed nonword options to fill a sentence gap. Each option utilized the same nonword base but with different real suffixes, only one of which was syntactically compatible with the sentence. The task included 24 items, with a maximum score of 24.

Relational knowledge was assessed using a morphological relatedness test adapted from Mahony et al. (2000). Participants were instructed to determine whether 20 word pairs were morphologically related, with a maximum score of 20.

Distributional knowledge was assessed using a task adapted from Study 1 in Friedline (2011). This test comprised 30 items that tested correct and incorrect ordering of derivational morphemes (e.g., *hopenessful* vs. *playfulness*). Participants were offered a six-degree scale to rate the “realness” of the words, with a maximum score of 30. The test materials for these assessments are provided in Appendix D. Participants’ language proficiency, morphological awareness, and related demographic information are summarized in Table 4.

Procedures

Potential participants completed an online background questionnaire through Qualtrics (www.qualtrics.com). Eligibility criteria included being non-English majors, Chinese-speaking learners of English, and postgraduate students. Qualified participants were then invited to the lab, where they completed a consent form, the Updated Vocabulary Levels Test (Webb et al., 2017), and the pre-test. Those who demonstrated mastery of the most frequent 2,000 English word families were invited back one week later to complete the eye-tracking portion of the study. Immediately following the eye movement data collection, an unexpected post-test was administered, after which the morphological awareness test and the LexTALE (Lemhöfer & Broersma, 2012) were administered.

Eye movement data were collected using an EyeLink Portable Duo eye tracker (SR Research Ltd.; <https://www.srresearch.com/>) at a sampling rate of 1,000 Hz. Participants were instructed to place their chins on a chin rest while seated 65 cm from the monitor. Stimuli were presented in Courier New font, size 24, and double-spaced, with each letter averaging 19 pixels on a 13-inch monitor, resulting in a visual angle of approximately 0.47 degrees. Participants read binocularly, but only movements of the right eye were tracked and recorded. The experiment consisted of 113 trials, including 1 practice trial, 64 stimulus trials, and 48 filler trials, evenly divided into three sections. Prior to each section, a 9-point calibration was

conducted, with short breaks provided between sections to allow participants to rest. A 1-point drift correction was performed before each trial to determine if further calibration was necessary.

Participants first viewed an introductory screen containing instructions indicating they would read English paragraphs on the screen, some followed by true/false comprehension questions. A practice trial familiarized participants with the experiment procedure. Targeted pseudowords were never positioned as the first or last word on a line. Participants could press the spacebar or respond to true/false questions using two designated keyboard buttons to advance. Participants could not access previous screens.

Each three-sentence reading context included three areas of interest (AOIs). The height of each AOI was set to 74 pixels, corresponding to the default double-spaced text in Eyelink software, while widths varied according to AOI length. The first AOI comprised the five-letter pseudo base word, measuring 7,696 pixels. The second AOI included the targeted suffix, measuring 4,884 pixels for three-letter suffixes and 6,290 pixels for four-letter suffixes. The third AOI encompassed the informative context following the pseudowords, with variable widths due to the context's length.

Data preparation and analysis

For the offline form recognition, meaning recall, and meaning recognition tests, incorrect responses were coded as 0 and correct responses as 1. The meaning recall test was independently scored by two research assistants. Interrater reliability for the pre-test: Cohen's kappa = 1.000, $p < .001$, given that there was only one correct response to score; post-test: Cohen's kappa = .954, $p < .001$. Raw eye-tracking data were pre-processed using Eyelink Data Viewer software (SR Research). Fixation patterns for each trial were reviewed, and vertical drift was manually adjusted as necessary. Fixations shorter than 80 ms were merged with adjacent fixations within 0.5 degrees of visual angle. Subsequently, fixations shorter than 80 ms or longer than 1200 ms were deleted. The eye fixation measures analyzed included total reading times for three AOIs: pseudo base words, targeted suffixes, and informative contexts.

Statistical analyses were conducted using R (R Core Team, 2022). To answer Research Question 1, two paired-samples *t*-tests were performed to compare pre-test and post-test scores on form recognition and meaning recognition. A Wilcoxon Signed-Rank Test was conducted for meaning recall due to the nonnormality of the data.

To address Research Question 2, three mixed-effect logistic regression models were built to identify factors predicting participants' accuracy on the three post-tests. These models were implemented using functions from the *lme4* package (version 1.1.21; Bates et al., 2015). For all models except the one for meaning recall, participants' pre-test responses were included as a covariate. Including pre-test scores for meaning recall resulted in a nearly singular Hessian, as participants had only one correct response on the pre-test, rendering this variable uninformative. Suffixes were coded as 0 for those with three letters and 1 for those with four letters. Continuous numerical variables were standardized as z-scores. These included total reading times for pseudo base words, targeted suffixes, and informative contexts; contextual variables such as richness ratings and length; participant scores on

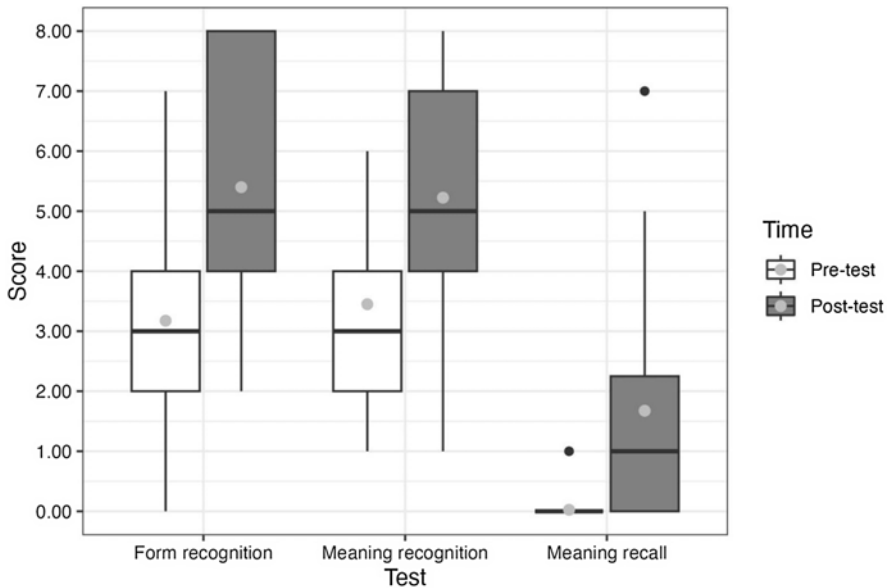


Figure 1. Suffix pre-test and post-test scores (max. score = 8). Bars within the box plots show medians and gray points show means.

vocabulary knowledge and morphological awareness tests, and suffix productivity and frequency. Informative context length, defined as the total number of letters, spaces, and punctuation marks, was treated as a control variable in the models.

Results

Suffix learning

To answer the first research question, we examined whether any of the three aspects of learners' suffix knowledge improved significantly after the treatment. Pre-test and post-test scores were calculated for each participant, and paired-samples *t*-tests and a Wilcoxon Signed-Rank Test were conducted to determine whether the observed increases were statistically significant. The alpha level was adjusted to 0.05/3 to account for the three comparisons (Al-Hoorie & Vitta, 2019). For form recognition, a paired-samples *t*-test showed that post-test scores ($M = 5.400$, $SD = 2.023$) were significantly higher than pre-test scores ($M = 3.175$, $SD = 1.647$), $t(39) = -6.022$, $p < .001$, Cohen's $d = 1.213$. Similarly, for meaning recognition, participants' accuracy in the post-test ($M = 5.225$, $SD = 1.941$) was significantly higher than their pre-test scores ($M = 3.450$, $SD = 1.280$), $t(39) = -5.211$, $p < .001$, Cohen's $d = 1.102$. For meaning recall, the Wilcoxon Signed-Rank Test indicated a significant improvement in post-test scores ($M = 1.675$, $SD = 1.789$) compared to pre-test scores ($M = 0.025$, $SD = 0.158$), $Z = -4.572$, $p < .001$, Cohen's $d = 1.686$. Cohen's d was calculated as Cohen's d_{av} for within-subject comparisons (Cumming, 2012). Figure 1 shows a comparison of pre-test and post-test scores for each aspect of suffix knowledge.

Table 5. Descriptive statistics for eye fixation measures on base word, suffix, and context

Variable (ms)	<i>M</i>	<i>SD</i>	95% CI
Total reading time (base word)	1723.591	1765.795	[1655.156, 1792.025]
Total reading time (suffix)	575.865	727.983	[547.651, 604.078]
Total reading time (context)	8378.302	4978.158	[8185.371, 8571.233]

Effects of predictor variables

Descriptive statistics for eye fixation measures, i.e., total reading times on pseudo base word, suffix, and context, are presented in Table 5.

To address the second research question, three multilevel logistic regression models were constructed, one for each suffix knowledge test: form recognition, meaning recall, and meaning recognition. These models examined the fixed effects of reading time fixation measures, learners' L2 proficiency (vocabulary knowledge and morphological awareness), and language-level factors (suffix length, suffix productivity, suffix frequency, and contextual richness) on the incidental learning of targeted suffixes.¹ The models included maximal random effects structures following the recommendations of Barr *et al.* (2013). However, the random effects required simplification to avoid either convergence issues or singular fits. Specifically, the form recognition model included a random intercept for participant and an uncorrelated random slope for morphological awareness. The meaning recall model included random intercepts for both participant and suffix items, while the meaning recognition model included only a random intercept for participant. The results of these three models are presented in Table 6.

For form recognition and meaning recall, total reading times on base words and suffixes positively predicted post-test response accuracy (form recognition: $b = 0.999$, $SE = 0.458$, $p = 0.029$, $OR = 2.715$ and $b = 0.774$, $SE = 0.360$, $p = 0.032$, $OR = 2.169$; meaning recall: $b = 1.328$, $SE = 0.405$, $p = .001$, $OR = 3.772$ and $b = 0.592$, $SE = 0.263$, $p = 0.024$, $OR = 1.808$). Odds ratios indicate that participants with reading times 1 SD longer than the mean were 2.715 and 2.169 times more likely to respond correctly on the form recognition post-test, and 3.772 and 1.808 times more likely on the meaning recall post-test, respectively. For meaning recognition, only total reading time on suffixes significantly predicted more accurate post-test responses ($b = 0.677$, $SE = 0.258$, $p = 0.009$, $OR = 1.968$).

In contrast, total reading time on contexts negatively impacted response accuracy across all post-tests (form recognition: $b = -0.794$, $SE = 0.310$, $p = 0.011$, $OR = 0.452$; meaning recall: $b = -1.359$, $SE = 0.384$, $p < .001$, $OR = 0.257$; and meaning recognition: $b = -0.908$, $SE = 0.229$, $p < .001$, $OR = 0.403$). For language-level factors, responses to 4-letter suffixes were significantly more accurate than 3-letter suffixes, particularly for form recognition ($b = 1.742$, $SE = 0.755$, $p = 0.021$, $OR = 5.706$). Suffix productivity also positively influenced meaning recall scores ($b = 1.811$, $SE = 0.598$, $p = 0.002$, $OR = 6.117$). However, contextual richness and suffix frequency did not significantly affect performance on any test. For language proficiency, better morphological awareness significantly improved meaning recall accuracy ($b = 0.517$, $SE = 0.250$, $p = 0.038$, $OR = 1.677$), while richer vocabulary

Table 6. Effects of eye fixation measures, participant’s language proficiency, and language-level factors on suffix knowledge learning gains

	Form Recognition					Meaning Recall					Meaning Recognition				
Fixed effects	<i>b</i>	<i>SE</i>	<i>z</i>	<i>p</i>	<i>OR</i>	<i>b</i>	<i>SE</i>	<i>z</i>	<i>p</i>	<i>OR</i>	<i>b</i>	<i>SE</i>	<i>z</i>	<i>p</i>	<i>OR</i>
Intercept	−0.058	0.423	−0.138	.890	0.943	−2.110	0.910	−2.319	.020	0.121	0.151	0.370	0.409	.683	1.163
Pre-test	1.129	0.336	3.363	<.001	3.093						0.660	0.281	2.348	.019	1.935
Vocabulary knowledge	0.375	0.224	1.673	.094	1.455	0.395	0.254	1.556	.120	1.485	0.349	0.175	1.997	.046	1.417
Morphological awareness	0.132	0.201	0.659	.510	1.142	0.517	0.250	2.070	.038	1.677	0.014	0.145	0.098	.922	1.014
Total reading time (base word)	0.999	0.458	2.178	.029	2.715	1.328	0.405	3.276	.001	3.772	0.481	0.260	1.852	.064	1.617
Total reading time (suffix)	0.774	0.360	2.149	.032	2.169	0.592	0.263	2.251	.024	1.808	0.677	0.258	2.620	.009	1.968
Total reading time (context)	−0.794	0.310	−2.557	.011	0.452	−1.359	0.384	−3.537	<.001	0.257	−0.908	0.229	−3.958	<.001	0.403
Contextual richness	0.453	0.274	1.652	.099	1.573	−1.018	0.594	−1.714	.087	0.361	−0.043	0.242	−0.180	.857	0.957
Suffix length	1.742	0.755	2.306	.021	5.706	−0.723	1.566	−0.462	.644	0.485	0.822	0.649	1.267	.205	2.274
Suffix productivity	0.108	0.289	0.372	.710	1.114	1.811	0.598	3.028	.002	6.117	0.081	0.236	0.345	.730	1.085
Suffix frequency	0.185	0.195	0.949	.342	1.203	−0.057	0.441	−0.128	.898	0.945	0.134	0.181	0.740	.459	1.143
<i>Control Variable</i>															
Context length	0.149	0.204	0.728	.466	1.161	0.318	0.447	0.712	.477	1.374	0.228	0.197	1.156	.248	1.256
Random effects	Variance					Variance					Variance				
by participant	Intercept	0.495			0.703	Intercept	0.668			0.817	Intercept	0.121			0.347
	Ma ^a	0.056			0.237										
by item						Intercept	0.460			0.678					

Note: Vocabulary knowledge, morphological awareness, total reading time (base word), total reading time (suffix), total reading time (context), contextual richness, suffix productivity, suffix frequency, and context length were all standardized as z scores. Total reading time (base word), total reading time (suffix), and total reading time (context) are given in ms. Values of *p* in bold face are significant at the .05 level of alpha.

^aMorphological awareness.

knowledge positively impacted meaning recognition scores ($b = 0.349$, $SE = 0.175$, $p = 0.046$, $OR = 1.417$).

Discussion

To investigate whether Chinese EFL learners could make incidental learning gains in three aspects of suffix knowledge, we examined pre-test and post-test accuracy for eight neutral derivational noun suffixes presented within three-sentence contexts. To gain a more comprehensive understanding of how learners' attention to morphological structure may contribute to learning gains, we also collected reading time data using eye tracking. These reading times, along with variables reflecting participants' L2 proficiency and language-level factors, were analyzed to identify predictors of accuracy on the three offline tests. Research Question 1 addressed the possibility of incidental learning of suffixes during reading, and the data showed that high-intermediate Chinese EFL learners could incidentally learn both form and meaning of suffixes through reading. Research Question 2 investigated the factors that potentially contributed to incidental learning across the three aspects of suffix knowledge (i.e., suffix form recognition, meaning recognition, and meaning recall). Our results indicated that attentional eye fixation measures, language-level features, and learners' language proficiency had an impact on learners' gains in suffix knowledge.

Suffix learning

Research Question 1 aimed to investigate whether participants could make learning gains through reading pseudowords presented within a context that provided insight into the meaning of each word and, subsequently, the meaning of the corresponding suffix. We compared pre-test and post-test scores for form recognition, meaning recall, and meaning recognition—three common measures of word knowledge used in incidental vocabulary learning research. Statistical analyses revealed significant improvements in the Chinese EFL learners' knowledge of both the form and meaning of the targeted suffixes after the reading session. Our findings regarding the incidental learning of L2 suffixes parallel findings on the incidental learning of L2 vocabulary in multiple ways. First, they suggest that L2 reading can also support the incidental learning of affixes, aligning with studies showing that reading can foster the incidental learning of L2 vocabulary.

Additionally, the significant learning gains observed in this study indicate that eight repeated exposures to the suffixes were sufficient for incidental learning. This finding aligns with previous L2 vocabulary studies, which have found that eight exposures to a new word are generally sufficient for incidental learning (Horst *et al.*, 1998; Waring & Takaki, 2003; Pellicer-Sánchez, 2016). However, unlike traditional vocabulary studies where learners encounter the same word multiple times, participants in this study encountered eight unique words, each sharing the same suffix. This design emphasizes the suffix itself rather than the entire pseudoword incorporating the suffix, more closely reflecting natural learning contexts where learners are likely to encounter different words that share a common suffix. The learning gains observed in this study for suffixes suggest that, like findings for words

in previous studies (e.g., Pellicer-Sánchez, 2016), the frequency threshold of eight exposures is enough for learning either. Moreover, the learning gains suggest that the high-intermediate Chinese EFL learners in this study successfully decomposed the morphologically complex words encountered during reading into pseudo base words and targeted suffixes. Thus, this study supports the view that intermediate to advanced L2 learners can effectively decompose morphologically complex words into their bases and derivational affixes (Diependaele et al., 2011; Li et al., 2017; Iwaizumi & Webb, 2023).

Similar to findings from previous research on incidental L2 vocabulary learning (Schmitt, 2008, 2010; Nation, 2013; Mohamed, 2018), this study found that among the three types of word knowledge, form recognition was the easiest for learners, followed by meaning recognition, and finally meaning recall. In this study, the greatest learning gains were observed in suffix form recognition, with an average increase of 2.22 correct answers on the eight-item post-test. This was followed by suffix meaning recognition, which showed an average increase of 1.78 correct answers, and suffix meaning recall, which had the smallest increase of 1.65 correct answers. These findings suggest that the learning difficulty pattern observed in vocabulary acquisition also applies to the acquisition of derivational suffixes.

Effects of predictor variables

Attentional fixation measures

The attentional fixation measures used in this study included total reading times on pseudo base words, targeted suffixes, and informative contexts. Two of these measures—total reading time on base words and suffixes—captured the time participants spent fixating on targeted pseudowords. Previous research on incidental vocabulary learning has consistently shown that total reading time on novel words significantly and positively predicts learning success for those words (Godfroid et al., 2013; Pellicer-Sánchez, 2016; Mohamed, 2018). In alignment with these findings, our study revealed that total reading time on targeted novel suffixes significantly and positively predicted learning gains across all three suffix knowledge tests. This positive relationship between attention to suffixes and suffix knowledge learning supports theoretical accounts of L2 acquisition, such as Schmidt's noticing hypothesis (1994, 1995, and 2001), which underscores the critical role of attention in language learning.

Additionally, our L2 learners' attention to both constituent morphemes (i.e., pseudo base words and suffixes) within the pseudowords provides insight into how morphological information was applied during sentential reading. The positive relationships between attention to suffixes and suffix learning gains, and between attention to base words and suffix learning gains, suggest that both suffixes and base words were effectively exploited as cues to support suffix learning when learners encountered unknown morphologically complex words. This finding contrasts with prior research that has tended to emphasize the prominent role of either component morpheme (i.e., *suffix effect* vs. *embedded stem activation account*; e.g., Beyersmann et al., 2015; Beyersmann et al., 2016; Grainger & Beyersmann, 2017; Amenta et al., 2025). In previous behavioral lexical decision paradigms, where (pseudo)words are presented in isolation, the relative contribution of each

morpheme has been investigated primarily through comparisons of participants' rejection accuracy and reaction times across different types of pseudowords (e.g., Casalis et al., 2015; Li et al., 2017; Amenta et al., 2025). Under such time-constrained conditions, participants are required to make rapid lexical decisions. For example, Amenta et al. (2025) reported that higher-proficiency participants rejected pseudowords with a real base word and a non-existent suffix more quickly than those with a non-existent base word and a real suffix (a condition that parallels our pseudoword design), suggesting a reliance on suffixes as morphological cues. Similarly, Casalis et al. (2015) and Amenta et al. (2025) found that words containing real suffixes were more likely to be recognized than those without, regardless of language proficiency. Research with native speakers has also demonstrated suffix effects through priming paradigms—for instance, Duñabeitia et al. (2008) used word pairs such as *baker*–*WALKER* to illustrate suffix priming. Regarding the role of base words in morphological processing and visual word recognition, studies have shown that low-proficiency learners are more likely to recognize words containing real base words or to reject pseudowords with non-existent bases (Amenta et al., 2025; Casalis et al., 2015). Such effects have also been documented among native speakers in priming experiments using pairs like *adaptable*–*ADAPTER* (Rastle et al., 2000). Collectively, these findings suggest that low-proficiency learners are particularly sensitive to embedded base words as morphological cues.

However, findings from lexical decision tasks may not fully reflect learners' comprehensive attention allocation or processing strategies when engaging with morphologically complex words. Rather, under time pressure, learners may prioritize the most accessible or diagnostic cues to make quick decisions. In contrast, our study employed a self-paced silent reading task within sentential contexts, where participants were likely motivated to derive meaning from the complex words in context. This may have encouraged dual attention to both suffixes and base words. The consistent appearance of a shared suffix across several novel words may have further prompted learners to recognize its form and infer its meaning. Unlike the case of novel-derived word processing, where attention to both morphemes is often attributed to morphemic interference (Taft & Forster, 1975), our findings suggest that attention was directed toward the integration of morphemes for meaning construction. Thus, attention to both suffixes and pseudo base words indicates that higher-proficiency L2 learners are sensitive to constituent morphemes even when encountering unknown words in context. This supports the view that morphological decomposition occurs during natural reading (Amenta et al., 2015) and is consistent with findings from priming and lexical decision tasks that have demonstrated morphological decomposition in L2 learners (Diependaele et al., 2011; Duñabeitia et al., 2013; Casalis et al., 2015; Zhang et al., 2016; Li et al., 2017; Amenta et al., 2025).

In addition, the relationship between total reading time on base words and suffix form recognition performance can potentially be explained by the phenomenon of parafoveal preview (Schotter et al., 2012), where readers gather information about upcoming words or morphemes while fixating on the current one. As demonstrated in compound words processing (Hyönä et al., 2004), parafoveal preview may also account for the relationship between attention to base words and suffix form recognition. This mechanism may also underlie why total reading time on base

words predicts meaning recall accuracy, suggesting that suffix meaning can be inferred from broader contextual engagement.

In contrast to the positive relationship observed between reading times on targeted words and learning outcomes, we found a negative relationship between reading times on contexts and all three measures of suffix knowledge. The underlying reasons for this finding are less well understood, as fixation-based measures on contextual regions have rarely been the focus of previous eye-tracking studies on incidental vocabulary learning. One possible explanation lies in the larger size of the context regions as AOIs. Fixation durations on these broader regions likely encompass not only attention to the contextual input but also the cognitive effort required to process and comprehend that input. This blend of attention and processing demands may help explain the observed negative correlation. Additionally, the role of language proficiency suggested by the influence of vocabulary knowledge provides a further possible interpretation: faster readers, who are often more proficient, may have required less time to extract relevant information from context and were thus better positioned to consolidate morphological knowledge. In this view, shorter reading times on context may reflect more efficient processing rather than reduced engagement.

Language-level factors

For the orthographic length of suffix, we found that long (i.e., four-letter) suffixes were more likely to be successfully learned than short (i.e., three-letter) suffixes in the form recognition test. This advantage may result from differences in phonological patterns between the two types of suffixes. Phonologically, a syllable typically consists of a vowel nucleus, an optional consonant onset preceding the vowel, and an optional consonant coda following it. All the long suffixes in our study began with a consonant onset, whereas three of the four targeted short suffixes began with a vowel nucleus. Consequently, the vowel nuclei in these short suffixes may have integrated with the consonant codas of their preceding pseudo bases' syllables when the pseudowords were pronounced. This integration likely made it more challenging for learners to decompose the pseudo base words from the nucleus-initial short suffixes during reading, adversely affecting learning gains for both form and meaning knowledge. Another factor influencing the learning of short suffixes could be their phonological properties. Three nucleus-initial short suffixes are considered non-neutral in some contexts (Fudge, 1984), potentially increasing their learning difficulty for L2 learners (Tyler & Nagy, 1989; Friedline, 2011). However, during the suffix selection phase, we followed Trevian's (2015) suggestion that examples exhibiting non-neutral properties in short suffixes are exceptional cases. Therefore, we classified these short suffixes as neutral. Nevertheless, the findings may reflect the "mixed" nature (i.e., sometimes neutral, sometimes non-neutral) of these short suffixes. An alternative explanation is that long suffixes, being perceptually salient (Laudanna & Burani, 1995; Bertram et al., 1999), prompt readers to adopt a decomposition strategy during morphological processing, where stems and suffixes are processed separately. This separate processing likely draws learners' attention to the long suffixes, resulting in improved recognition of their forms (Kuperman et al., 2010). Amenta et al. (2015)

provided additional insight using eye fixation data, observing that the first fixation location in derived words typically falls on the stem. Long suffixes, positioned in the periphery of vision, may experience reduced visual acuity and limited visibility, which can hinder comprehensive morphological and early semantic processing. However, this incomplete early processing may redirect readers' attention specifically to the long suffixes, thereby facilitating their form recognition.

In relation to the two key suffix properties relevant to morphological processing, previous research involving both L1 speakers (e.g., Bertram *et al.*, 1999; Bertram *et al.*, 2000; Lázaro *et al.*, 2015; Sánchez-Gutiérrez *et al.*, 2018) and L2 learners (e.g., Burani & Thornton, 2003; Silva & Clahsen, 2008; Dal Maso & Giraudo, 2014) has shown facilitatory effects of affix productivity and affix frequency through various experimental paradigms. Specifically, more productive and more frequent affixes have been associated with enhanced word recognition and, in some cases, greater difficulty in rejecting pseudowords. In our study, suffix productivity significantly predicted L2 learners' success in meaning recall, suggesting that production-based task is particularly sensitive to variation at the suffix level. In contrast, suffix frequency did not predict any learning outcomes. These findings highlight the beneficial role of suffix productivity, especially in the most cognitively demanding dimension of suffix knowledge: meaning recall. Notably, suffix productivity only predicted performance on the production-based suffix knowledge task. This pattern aligns with prior findings from L1 English-speaking children, who used highly productive suffixes more effectively than less productive ones in elicited production tasks (Windsor & Hwang, 1999). Conversely, Schmidtke *et al.* (2022) found that L2 learners made greater gains with less productive suffixes in a written suffix production task. However, their participants received eight months of language exposure, whereas learners in our study had a significantly shorter learning duration. The discrepancy likely reflects differences in exposure time, which may modulate the role of suffix productivity during learning. Regarding suffix frequency, its lack of predictive value in our study may stem from our experimental design. During the selection process, suffixes already recognized by most pilot participants were excluded, ensuring that target suffixes were low in frequency. Consequently, the limited variability in suffix frequency likely reduced its potential impact on learning outcomes.

Similarly, contextual richness did not have a significant impact on any suffix post-tests. This was somewhat unexpected, as contextual richness is often associated with gains in word meaning knowledge (Webb, 2008). However, this result may be explained by the design of our study. Using a 1-to-3 rating scale, all contexts were informative as intended, with an average rating of 2.31 ($SD = 0.08$, range = 2.18–2.38; see Table 3). The lack of variation in contextual richness within the study may account for its limited influence on suffix knowledge gains. Notably, previous eye-tracking studies on incidental vocabulary learning have not examined the contextual richness of individual reading contexts (e.g., Godfroid *et al.*, 2013; Pellicer-Sánchez, 2016; Mohamed, 2018). Future research with a broader range of contextual richness will be necessary to clarify its potential relationship with the learning of the three aspects of suffix knowledge.

Learner's language proficiency

Our results provide evidence that learners' vocabulary knowledge predicts incidental learning gains in suffix knowledge, though this effect was observed in only one of the three post-tests. Similar to previous research showing a positive relationship between language proficiency and word meaning recognition (e.g., Tekmen & Daloğlu, 2006), our research found that language proficiency, as measured by vocabulary knowledge, positively and significantly predicted suffix meaning recognition success. Language proficiency has been discussed as a key factor influencing L2 learners' morphological processing strategies. Highly proficient learners are more likely to decompose morphologically complex words, similar to native speakers, whereas less proficient learners tend to rely on whole word recognition (e.g., Silva & Clahsen, 2008; Li et al., 2017). The learning gains in suffix meaning recognition observed among our high intermediate proficiency learners support the decomposition-based strategy, as successful parsing of complex words into base words and suffixes is a prerequisite for understanding suffix learning. However, it is noteworthy that language proficiency predicted gains only in suffix meaning knowledge, not in form-related knowledge. This selective effect may be attributed to the nature of our experimental context, which involved reading for comprehension, a task that naturally emphasizes semantic processing over form-focused learning.

Nevertheless, the relationship between vocabulary knowledge and suffix learning gains should not be considered unidirectional or entirely direct. As with findings from research on the correlation between affix knowledge and L2 vocabulary (Mochizuki & Aizawa, 2000), our results indicate that the learning gains in affixes and vocabulary knowledge likely have a mutually beneficial relationship. Additionally, vocabulary knowledge, as a measure of language proficiency, is strongly correlated with learners' reading comprehension (e.g., Laufer, 1992). Improvements in reading comprehension can enhance learners' ability to understand sentence contexts, thereby facilitating the inference and learning of an unknown suffix's meaning. Thus, reading comprehension may potentially mediate the relationship between suffix meaning learning gains and L2 vocabulary knowledge. Interestingly, vocabulary knowledge only significantly predicted suffix meaning recognition, not suffix meaning recall. This discrepancy may be attributed to the greater task difficulty associated with suffix meaning recall, which may require higher-frequency encounters with a given suffix for L2 learners to pick up some productive knowledge of it.

Morphological awareness for derivations only positively and significantly predicted learning gains in suffix meaning recall. Morphological awareness works when learners encounter morphologically complex words during reading, enabling them to recognize the morphemic structure of these words (Carlisle, 1995). However, this recognition of morphemic structure provides access only to the components of the word and may not necessarily reflect whether learners have fully learned the suffixes. Nevertheless, our study found a positive and significant relationship between morphological awareness and suffix meaning recall. As previously noted, the suffix meaning recall test is more difficult than the other two suffix knowledge tests since it measures productive knowledge, which requires test takers to provide the meaning of the suffix. The significant relationship between

morphological awareness and learners' performance on this more demanding test may indicate that suffix knowledge recall draws on learners' morphological awareness to a greater degree than recognition does.

Limitations and implications

The first limitation of the present study is the inability to fully control for participants' prior knowledge of the targeted suffixes. Although we used piloting to identify low-frequency suffixes that participants were unlikely to know, it is still possible that some participants had prior knowledge or partial familiarity with these suffixes. Another limitation lies in the ambiguity regarding whether the length of the suffix or another factor, such as the potential mixed neutral/non-neutral distinction, explains the greater incidental learning observed for four-letter suffixes. Finally, the study does not provide insight into whether participants will retain the knowledge demonstrated in the post-tests, as no delayed post-test was included to assess retention. Thus, future research should incorporate delayed post-tests to evaluate both suffix learning and retention. Additionally, future studies could benefit from incorporating derivational suffixes that form words of other parts of speech, non-neutral suffixes, and/or prefixes.

For L2 teaching, the findings of the present study align with previous research suggesting that reading in a L2 is beneficial for language learning. Furthermore, our study demonstrates that reading in an L2 supports the learning of unfamiliar suffixes in a manner comparable to its effect on learning mid-to-low frequency words. Thus, incidental learning during reading can serve as a complement to explicit teaching, particularly for high-frequency suffixes. Based on these findings, educational institutions and publishers may consider developing graded readers that systematically enhance learners' suffix knowledge, following the established affix difficulty order (Bauer & Nation, 1993; Sasao & Webb, 2017). However, as there are few studies on the magnitude of incidental suffix learning gains, it is challenging to make direct comparisons or draw definitive conclusions. Despite this, incidental learning can still be considered a valuable complement to the explicit teaching of high-frequency affixes, much like its role in vocabulary acquisition (e.g., Schmitt & Carter, 2000; Wu, 2009; Hu, 2013).

Conclusion

This study builds upon previous research on incidental vocabulary acquisition through reading by specifically examining English neutral derivational noun suffixes. The significant improvements in accuracy on the offline post-tests indicate that Chinese EFL learners can effectively acquire both the form and meaning of suffixes through incidental learning during reading. These findings highlight that L2 readers are capable of developing sensitivity to the morphological structures of a second language. Furthermore, incidental learning while reading proves to be a valuable supplement to intentional learning, particularly for English affixes that occur at lower frequencies. Additionally, the investigation into attentional fixation measures, learners' language proficiency, and language-level features demonstrates that these factors play a role in influencing suffix learning gains.

Replication package. Data availability: the data that support the findings of this study are openly available in Open Science Framework at <https://osf.io/xz49r/>.

Acknowledgements. This research was supported by the University of Macau under research grant numbers MYRG-GRG2023-00112-FED.

Note

1 We thank the reviewer for raising the issue of potential collinearity among the fixed-effect predictors. In response, we conducted diagnostics using the `check_collinearity()` function from the *performance* R package. Variance inflation factors (VIFs) for *Suffix Productivity* and *Suffix Frequency* ranged from 1.90 to 3.63 across models, all well below the conventional threshold of 5. Additionally, the Pearson correlation between the two predictors was $r = -0.27$, indicating only modest collinearity.

References

- Al-Hoorie, A. H., & Vitta, J. P. (2019). The seven sins of L2 research: a review of 30 journals' statistical quality and their CiteScore, SJR, SNIP, JCR Impact Factors. *Language Teaching Research*, 23(6), 727–744. <https://doi.org/10.1177/1362168818767191>
- Amenta, S., Foppolo, F., & Badan, L. (2025). The role of morphological information in processing pseudo-words in Italian L2 learners: it's a matter of experience. *Journal of Cognition*, 8(1), 1–17. <https://doi.org/10.5334/joc.420>
- Amenta, S., Marelli, M., & Crepaldi, D. (2015). The fruitless effort of growing a fruitless tree: Early morpho-orthographic and morpho-semantic effects in sentence reading. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 41(5), 1587–1596. <https://doi.org/10.1037/xlm0000104>
- Baayen, R. H. (1994). Productivity in language production. *Language and Cognitive Processes*, 9(3), 447–469. <https://doi.org/10.1080/01690969408402127>
- Barr, D. J., Levy, R., Scheepers, C., & Tily, H. J. (2013). Random effects structure for confirmatory hypothesis testing: keep it maximal. *Journal of Memory and Language*, 68(3), 255–278. <https://doi.org/10.1016/j.jml.2012.11.001>
- Bates D., Mächler M., Bolker B., & Walker S. (2015). Fitting linear mixed-effects models using lme4. *Journal of Statistical Software*, 67(48), 1–48. <https://doi.org/10.18637/jss.v067.i01>.
- Bauer, L., & Nation, P. (1993). Word families. *International Journal of Lexicography*, 6(4), 253–279. <https://doi.org/10.1093/ijl/6.4.253>
- Beauvillain, C. (1996). The Integration of morphological and whole-word form information during eye fixations on prefixed and suffixed words. *Journal of Memory and Language*, 35(6), 801–820. <https://doi.org/10.1006/jmla.1996.0041>
- Bertram, R., Laine, M., & Karvinen, K. (1999). The interplay of word formation type, affixal homonymy, and productivity in lexical processing: evidence from a morphologically rich language. *Journal of Psycholinguistic Research*, 28, 213–226. <https://doi.org/10.1023/A:1023200313787>
- Bertram, R., Schreuder, R., & Baayen, R. H. (2000). The balance of storage and computation in morphological processing: the role of word formation type, affixal homonymy, and productivity. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 26(2), 489–511. <https://doi.org/10.1037/0278-7393.26.2.489>
- Beyersmann, E., Castles, A., & Coltheart, M. (2012). Morphological processing during visual word recognition in developing readers: evidence from masked priming. *The Quarterly Journal of Experimental Psychology*, 65(7), 1306–1326. <https://doi.org/10.1080/17470218.2012.656661>
- Beyersmann, E., Casalis, S., Ziegler, J. C., & Grainger, J. (2015). Language proficiency and morpho-orthographic segmentation. *Psychonomic Bulletin & Review*, 22, 1054–1061. <https://doi.org/10.3758/s13423-014-0752-9>
- Beyersmann, E., Cavalli, E., Casalis, S., & Colé, P. (2016). Embedded stem priming effects in prefixed and suffixed pseudowords. *Scientific Studies of Reading*, 20(3), 220–230. <https://doi.org/10.1080/10888438.2016.1140769>

- Burani, C. & Thornton, A. (2003). The Interplay of Root, Suffix and Whole-Word Frequency in Processing Derived Words. In R. Baayen & R. Schreuder (Eds.), *Morphological structure in language processing* (pp. 157–208). Berlin, New York: De Gruyter Mouton. <https://doi.org/10.1515/9783110910186.157>
- Carlisle, J. F. (1995). Morphological awareness and early reading achievement. In L. B. Feldman (Ed.), *Morphological Aspects of Language Processing* (pp. 189–209). Lawrence Erlbaum Associates, Inc.
- Casalis, S., Comissaire, E., & Duncan, L. G. (2015). Sensitivity to morpheme units in English as L2 word recognition. *Writing Systems Research*, 7(2), 186–201. <https://doi.org/10.1080/17586801.2014.976165>
- Chen, C., & Truscott, J. (2010). The effects of repetition and L1 lexicalization on incidental vocabulary acquisition. *Applied Linguistics*, 31(5), 693–713. <https://doi.org/10.1093/applin/amq031>
- Cobb, T. (n.d.). Compleat Lexical Tutor (Version 8). Available online at <http://www.lexutor.ca/> (accessed August 20, 2022).
- Conklin, K., & Pellicer-Sánchez, A. (2016). Using eye-tracking in applied linguistics and second language research. *Second Language Research*, 32(3), 453–467. <https://doi.org/10.1177/0267658316637401>
- Council of Europe. (2001). *Common European Framework of Reference for Languages: Learning, Teaching, Assessment*. Cambridge University Press.
- Cumming, G. (2012). *Understanding the new statistics: effect sizes, confidence intervals, and meta-analysis*. Routledge.
- Dal Maso, S., & Giraudo, H. (2014). Morphological processing in L2 Italian: Evidence from a masked priming study. *Linguisticae Investigationes*, 37(2), 322–337. <https://doi.org/10.1075/li.37.2.09mas>
- Daneman, M., & Carpenter, P. A. (1983). Individual differences in integrating information between and within sentences. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 9(4), 561–584. <https://doi.org/10.1037/0278-7393.9.4.561>
- de Varda, A. G., Gatti, D., Marelli, M., & Günther, F. (2024). Meaning beyond lexicality: capturing pseudoword definitions with language models. *Computational Linguistics*, 50(4), 1313–1343. https://doi.org/10.1162/coli_a_00527
- Diependaele, K., Duñabeitia, J. A., Morris, J., & Keuleers, E. (2011). Fast morphological effects in first and second language word recognition. *Journal of Memory and Language*, 64(4), 344–358. <https://doi.org/10.1016/j.jml.2011.01.003>
- Duñabeitia, J. A., Dimitropoulou, M., Morris, J., & Diependaele, K. (2013). The role of form in morphological priming: evidence from bilinguals. *Language and Cognitive Processes*, 28(7), 967–987. <https://doi.org/10.1080/01690965.2012.713972>
- Duñabeitia, J. A., Perea, M., & Carreiras, M. (2008). Does darkness lead to happiness? Masked suffix priming effects. *Language and Cognitive Processes*, 23(7–8), 1002–1020. <https://doi.org/10.1080/01690960802164242>
- Elgort, I., Brysbaert, M., Stevens, M., & van Assche, E. (2018). Contextual word learning during reading in a second language: an eye-movement study. *Studies in Second Language Acquisition*, 40(2), 341–366. <https://doi.org/10.1017/S0272263117000109>
- Elgort, I., & Warren, P. (2014). L2 vocabulary learning from reading: Explicit and tacit lexical knowledge and the role of learner and item variables. *Language Learning*, 64(2), 365–414. <https://doi.org/10.1111/lang.12052>
- Ellis, R. (1999). *Learning a second language through interaction*. John Benjamins.
- Friedline, B. E. (2011). Challenges in the second language acquisition of derivational morphology: from theory to practice. [Doctoral dissertation, University of Pittsburgh]. <https://d-scholarship.pitt.edu/8351/>
- Fudge, E. (1984). *English Word-Stress*. Allen & Unwin.
- Gatti, D., Marelli, M., & Rinaldi, L. (2023). Out-of-vocabulary but not meaningless: evidence for semantic-priming effects in pseudoword processing. *Journal of Experimental Psychology: General*, 152(3), 851–863. <https://doi.org/10.1037/xge0001304>
- Godfroid, A., Ahn, J., Choi, I., Ballard, L., Cui, Y., Johnston, S., Lee, S., Sarkar, A., & Yoon, H.-J. (2018). Incidental vocabulary learning in a natural reading context: an eye-tracking study. *Bilingualism: Language and Cognition*, 21(3), 563–584. <https://doi.org/10.1017/S1366728917000219>
- Godfroid, A., Boers, F., & Housen, A. (2013). An eye for words: Gauging the role of attention in incidental L2 vocabulary acquisition by means of eye-tracking. *Studies in Second Language Acquisition*, 35(3), 483–517. <https://doi.org/10.1017/S0272263113000119>

- Grainger, J., & Beyersman, E. (2017). Edge-aligned embedded word activation initiates morpho-orthographic segmentation. In B. H. Ross (Ed.), *The Psychology of Learning and Motivation* (pp. 285–317). Elsevier Academic Press.
- Hay, J. (2002). From speech perception to morphology: affix ordering revisited. *Language*, 78(3), 527–555. <https://doi.org/10.1353/lan.2002.0159>
- Hay, J. (2003). *Causes and Consequences of Word Structure*. Routledge.
- Horst, M., Cobb, T., & Meara, P. (1998). Beyond a clockwork orange: acquiring second language vocabulary through reading. *Reading in a Foreign Language*, 11(2), 207–223.
- Hu, H. M. (2013). The effects of word frequency and contextual types on vocabulary acquisition from extensive reading: a case study. *Journal of Language Teaching and Research*, 4(3), 487–495. <https://doi.org/10.4304/jltr.4.3.487-495>
- Hulstijn, J. H. (2001). Intentional and incidental second language vocabulary learning: A reappraisal of elaboration, rehearsal and automaticity. In P. Robinson (Ed.), *Cognition and Second Language Instruction* (pp. 258–286). Cambridge: Cambridge University Press.
- Hulstijn, J. H., Hollander, M., & Greidanus, T. (1996). Incidental vocabulary learning by advanced foreign language students: the influence of marginal glosses, dictionary use, and reoccurrence of unknown words. *The Modern Language Journal*, 80(3), 327–339. <https://doi.org/10.1111/j.1540-4781.1996.tb01614.x>
- Hyönä, J., Bertram, R., & Pollatsek, A. (2004). Are long compound words identified serially via their constituents? Evidence from an eye-movement-contingent display change study. *Memory & Cognition*, 32(4), 523–532. <https://doi.org/10.3758/BF03195844>
- Iwaizumi, E., & Webb, S. (2023). To what extent do learner- and word-related variables affect production of derivatives?. *Language Learning*, 73(1), 301–336. <https://doi.org/10.1111/lang.12524>
- Johns, B. T., Dye, M., & Jones, M. N. (2016). The influence of contextual diversity on word learning. *Psychonomic Bulletin & Review*, 23(4), 1214–1220. <https://doi.org/10.3758/s13423-015-0980-7>
- Joseph, H., & Nation, K. (2018). Examining incidental word learning during reading in children: the role of context. *Journal of Experimental Child Psychology*, 166, 190–211. <https://doi.org/10.1016/j.jecp.2017.08.010>
- Keuleers, E., & Brysbaert, M. (2010). Wuggy: a multilingual pseudoword generator. *Behavior Research Methods*, 42, 627–633. <https://doi.org/10.3758/BRM.42.3.627>
- Kraut, R. (2015). The relationship between morphological awareness and morphological decomposition among English language learners. *Reading and Writing*, 28(6), 873–890. <https://doi.org/10.1007/s1145-015-9553-4>
- Kuperman, V., Bertram, R., & Baayen, R. H. (2010). Processing trade-offs in the reading of Dutch derived words. *Journal of Memory and Language*, 62(2), 83–97. <https://doi.org/10.1016/j.jml.2009.10.001>
- Kwantes, P. J., & Mewhort, D. J. K. (1999). Evidence for sequential processing in visual word recognition. *Journal of Experimental Psychology: Human Perception and Performance*, 25(2), 376–381. <https://doi.org/10.1037/0096-1523.25.2.376>
- Laudanna, A., & Burani, C. (1995). Distributional Properties of Derivational Affixes: Implications for Processing. In L. B. Feldman (Ed.), *Morphological aspects of language processing* (pp. 345–364). Lawrence Erlbaum Associates.
- Laufer, B. (1992). How Much Lexis is Necessary for Reading Comprehension?, In P. J. L. Arnaud & H. Béjoint (Eds.), *Vocabulary and Applied Linguistics* (pp. 126–132). Palgrave Macmillan UK.
- Laufer, B. (2023). Understanding L2-derived words in context: Is complete receptive morphological knowledge necessary? *Studies in Second Language Acquisition*, 1–14. <https://doi.org/10.1017/s0272263123000219>
- Laufer, B., Webb, S., Kim, S.K., & Yohanan, B. (2021). How well do learners know derived words in a second language. The effect of proficiency, word frequency and type of affix. *ITL-International Journal of Applied Linguistics*, 172(2), 229258. <https://doi.org/10.1075/itl.20020.lau>
- Lázaro, M., Sainz, J., & Illera, V. (2015). The role of derivative suffix productivity in the visual word recognition of complex words. *Psicologica: International Journal of Methodology and Experimental Psychology*, 36(2), 165–184. <https://hdl.handle.net/20.500.14352/115338>
- Lemhöfer, K., & Broersma, M. (2012). Introducing LexTALE: a quick and valid lexical test for advanced learners of English. *Behavior Research Methods*, 44(2), 325–343. <https://doi.org/10.3758/s13428-011-0146-0>

- Li, J., Taft, M., & Xu, J. (2017). The processing of English derived words by Chinese-English bilinguals. *Language Learning*, 67(4), 858–884. <https://doi.org/10.1111/lang.12247>
- Lowell, R., & Morris, R. K. (2017). Impact of contextual constraint on vocabulary acquisition in reading. *Journal of Cognitive Psychology*, 29(5), 551–569. <https://doi.org/10.1080/20445911.2017.1299155>
- Mahony, D. L. (1994). Using sensitivity to word structure to explain variance in high school and college level reading ability. *Reading and Writing*, 6(1), 19–44. <https://doi.org/10.1007/BF01027276>
- Mahony, D., Singson, M., & Mann, V. (2000). Reading ability and sensitivity to morphological relations. *Reading and Writing*, 12(3), 191–218. <https://doi.org/10.1023/A:1008136012492>
- Medler, D.A., & Binder, J.R. (2005). MCWord: An On-Line Orthographic Database of the English Language. Available online at: <http://www.neuro.mcw.edu/mcword/> (accessed August 20, 2022).
- Mochizuki, M., & Aizawa, K. (2000). An affix acquisition order for EFL learners: an exploratory study. *System*, 28(2), 291–304. [https://doi.org/10.1016/S0346-251X\(00\)00013-0](https://doi.org/10.1016/S0346-251X(00)00013-0)
- Mohamed, A. (2018). Exposure frequency in L2 reading: an eye-movement perspective of incidental vocabulary learning. *Studies in Second Language Acquisition*, 40(2), 269–293. <https://doi.org/10.1017/S0272263117000092>
- Mousikou, P., & Schroeder, S. (2019). Morphological processing in single-word and sentence reading. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 45(5), 881–903. <https://doi.org/10.1037/xlm0000619>
- Nagy, W. E., Anderson, R. C., & Herman, P. A. (1987). Learning word meanings from context during normal reading. *American Educational Research Journal*, 24(2), 237–270. <https://doi.org/10.2307/1162893>
- Nation, I. S. (2012). The BNC/COCA word family lists. Available online at: <http://www.victoria.ac.nz/lals/about/staff/paul-nation> (accessed August 20, 2022)
- Nation, I. S. P. (2013). *Learning Vocabulary in Another Language*. Cambridge University Press.
- Nation, I. S. P. (2022). *Learning Vocabulary in Another Language*. Cambridge University Press.
- Pellicer-Sánchez, A. (2016). Incidental L2 vocabulary acquisition from and while reading: an eye-tracking study. *Studies in Second Language Acquisition*, 38(1), 97–130. <https://doi.org/10.1017/S0272263115000224>
- Pellicer-Sánchez, A., Conklin, K., & Vilkaitė-Lozdienė, L. (2021). The effect of pre-reading instruction on vocabulary learning: an investigation of L1 and L2 readers' eye movements. *Language Learning*, 71(1), 162–203. <https://doi.org/10.1111/lang.12430>
- Pellicer-Sánchez, A., & Schmitt, N. (2010). Incidental vocabulary acquisition from an authentic novel: Do “things fall apart”? *Reading in a Foreign Language*, 22(1), 31–55.
- R Core Team. (2022). *R: A language and environment for statistical computing (Version 4.1.3)*. R Foundation for Statistical Computing.
- Rastle, K., & Davis, M. H. (2008). Morphological decomposition based on the analysis of orthography. *Language and Cognitive Processes*, 23(7–8), 942–971. <https://doi.org/10.1080/01690960802069730>
- Rastle, K., Davis, M. H., Marslen-Wilson, W. D., & Tyler, L. K. (2000). Morphological and semantic effects in visual word recognition: a time-course study. *Language and Cognitive Processes*, 15(4–5), 507–537. <https://doi.org/10.1080/01690960050119689>
- Rastle, K., Harrington, J., & Coltheart, M. (2002). 358,534 nonwords: The ARC nonword database. *The Quarterly Journal of Experimental Psychology Section A*, 55(4), 1339–1362. <https://doi.org/10.1080/02724980244000099>
- 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 and Daloğlu (2006). *International Review of Applied Linguistics in Language Teaching*, 58(1), 75–102. <https://doi.org/10.1515/iral-2015-0115>
- Sánchez-Gutiérrez, C. H., Mailhot, H., Deacon, S. H., & Wilson, M. A. (2018). MorphoLex: a derivational morphological database for 70,000 English words. *Behavior Research Methods*, 50, 1568–1580. <https://doi.org/10.3758/s13428-017-0981-8>
- Sasao, Y., & Webb, S. (2017). The word part levels test. *Language Teaching Research*, 21(1), 12–30. <https://doi.org/10.1177/1362168815586083>
- Schmidtke, D., Rahmadian, S., & Moro, A. L. (2022). Morphological knowledge in English learner university students is sensitive to language statistics: a longitudinal study. *Applied Psycholinguistics*, 43(4), 889–919. <https://doi.org/10.1017/S0142716422000182>

- Schmitt, N. (2008). Instructed second language vocabulary learning. *Language Teaching Research*, 12(3), 329–363. <https://doi.org/10.1177/1362168808089921>
- Schmitt, N. (2010). *Researching Vocabulary: A Vocabulary Research Manual*. Palgrave Macmillan.
- Schmitt, N., & Carter, R. (2000). The lexical advantages of narrow reading for second language learners. *TESOL Journal*, 9(1), 4–9. <https://doi.org/10.1002/j.1949-3533.2000.tb00220.x>
- Schmitt, N., & Meara, P. (1997). Researching vocabulary through a word knowledge framework: Word associations and verbal suffixes. *Studies in Second Language Acquisition*, 19(1), 17–36. <https://doi.org/10.1017/S0272263197001022>
- Schmitt, N., & Zimmerman, C. B. (2002). Derivative word forms: what do learners know? *TESOL Quarterly*, 36(2), 145–171. <https://doi.org/10.2307/3588328>
- Schotter, E. R., Angele, B., & Rayner, K. (2012). Parafoveal processing in reading. *Attention, Perception, & Psychophysics*, 74(1), 5–35. <https://doi.org/10.3758/s13414-011-0219-2>
- Silva, R., & Clahsen, H. (2008). Morphologically complex words in L1 and L2 processing: evidence from masked priming experiments in English. *Bilingualism: Language and Cognition*, 11(2), 245–260. <https://doi.org/10.1017/S1366728908003404>
- Snoder, P., & Laufer, B. (2022). EFL learners' receptive knowledge of derived words: the case of Swedish adolescents. *TESOL Quarterly*, 56(4), 1242–1265. <https://doi.org/10.1002/tesq.3101>
- Taft, M., & Forster, K. I. (1975). Lexical storage and retrieval of prefixed words. *Journal of Verbal Learning and Verbal Behavior*, 14(6), 638–647. [https://doi.org/10.1016/S0022-5371\(75\)80051-X](https://doi.org/10.1016/S0022-5371(75)80051-X)
- Tekmen, E. A. F., & Daloğlu, A. (2006). An investigation of incidental vocabulary acquisition in relation to learner proficiency level and word frequency. *Foreign Language Annals*, 39(2), 220–243. <https://doi.org/10.1111/j.1944-9720.2006.tb02263.x>
- Trevian, I. (2015). *English Suffixes: Stress-assignment Properties, Productivity, Selection and Combinatorial Processes*. Peter Lang AG.
- Tyler, A., & Nagy, W. (1989). The acquisition of English derivational morphology. *Journal of Memory and Language*, 28(6), 649–667. [https://doi.org/10.1016/0749-596X\(89\)90002-8](https://doi.org/10.1016/0749-596X(89)90002-8)
- Uchihara, T., Webb, S., & Yanagisawa, A. (2019). The effects of repetition on incidental vocabulary learning: a meta-analysis of correlational studies. *Language Learning*, 69(3), 559–599. <https://doi.org/10.1111/lang.12343>
- Waring, R., & Takaki, M. (2003). At what rate do learners learn and retain new vocabulary from reading a graded reader? *Reading in a Foreign Language*, 15(2), 130–163.
- Webb, S. (2007). The effects of repetition on vocabulary knowledge. *Applied Linguistics*, 28(1), 46–65. <https://doi.org/10.1093/applin/aml048>
- Webb, S. (2008). The effects of context on incidental vocabulary learning. *Reading in a Foreign Language*, 20(2), 232–245.
- Webb, S. (2020). Incidental Vocabulary Learning. In Webb, S. (Ed.), *The routledge handbook of vocabulary studies* (pp. 225–239). Routledge.
- Webb, S., Sasao, Y., & Ballance, O. (2017). The updated vocabulary levels test: developing and validating two new forms of the VLT. *ITL - International Journal of Applied Linguistics*, 168(1), 33–69. <https://doi.org/10.1075/itl.168.1.02web>
- Webb, S., Uchihara, T., & Yanagisawa, A. (2023). How effective is second language incidental vocabulary learning? A meta-analysis. *Language Teaching*, 56(2), 1–20. <https://doi.org/10.1017/s0261444822000507>
- Windsor, J., & Hwang, M. (1999). Derivational suffix productivity for students with and without language-learning disabilities. *Journal of Speech, Language, and Hearing Research*, 42(1), 220–230. <https://doi.org/10.1044/jslhr.4201.220>
- Wu, L. (2009). *Incidental English Vocabulary Studying in L2 learning: A Study of Learning and Teaching English Vocabulary in a College in China*. (Master thesis, Kristianstad University College)
- Yarkoni, T., Balota, D., & Yap, M. (2008). Moving beyond Coltheart's N: a new measure of orthographic similarity. *Psychonomic Bulletin & Review*, 15(5), 971–979. <https://doi.org/10.3758/PBR.15.5.971>
- Zahar, R., Cobb, T., & Spada, N. (2001). Acquiring vocabulary through reading: effects of frequency and contextual richness. *The Canadian Modern Language Review*, 57(4), 541–572. <https://doi.org/10.3138/cmlr.57.4.541>

- Zhang, Q., Liang, L., Yao, P., Hu, S., & Chen, B.** (2016). Parallel morpho-orthographic and morpho-semantic activation in processing second language morphologically complex words: Evidence from Chinese-English bilinguals. *International Journal of Bilingualism*, **21**(3), 291–305. <https://doi.org/10.1177/13670069156242>
- Ziegler, J. C., Besson, M., Jacobs, A. M., Nazir, T. A., & Carr, T. H.** (1997). Word, pseudoword, and nonword processing: a multitask comparison using event-related brain potentials. *Journal of Cognitive Neuroscience*, **9**(6), 758–775. <https://doi.org/10.1162/jocn.1997.9.6.758>