

ORIGINAL ARTICLE

The effect of frequency of exposure on the processing and learning of collocations: A comparison of first and second language readers' eye movements

Ana Pellicer-Sánchez^{1,*} , Anna Siyanova-Chanturia^{2,3}  and Fabio Parente⁴ 

¹University College London, London, UK, ²Victoria University of Wellington, Wellington, New Zealand,

³Ocean University of China, Qingdao, China and ⁴University of Derby, Derby, UK

*Corresponding author. Email: a.pellicer-sanchez@ucl.ac.uk

(Received 12 February 2021; revised 1 March 2022; accepted 21 March 2022; first published online 04 May 2022)

Abstract

This study examined the processing and acquisition of novel words and their collocates (i.e., words that frequently co-occur with other words) from reading and the effect of frequency of exposure on this process. First and second language speakers of English read a story with 1) eight exposures of adjective-pseudoword collocations, 2) four exposures of the same collocations, or 3) eight exposures of control collocations. Results of recall and recognition tests showed that participants acquired knowledge not only of the form and meaning of the pseudowords but also of their collocates. The analysis of eye movements showed a significant effect of exposure on the processing of novel collocations for both first and second language readers, with reading times decreasing as a function of exposure. Eight exposures to novel adjective-pseudoword collocations were enough to develop processing speed comparable to that of known collocations. However, when analyzing the processing of the individual components of the collocations, results showed that eight exposures to the pseudowords were not enough for second language readers to develop processing speed comparable to known words. The frequency manipulation in the present study (four vs. eight exposures) did not lead to differences in the learning or processing of collocations. Finally, reading times were not a significant predictor of vocabulary gains.

Keywords: incidental learning; frequency of exposure; collocations; reading; eye movements

Introduction

Exposure is a key to second language (L2) incidental vocabulary learning (Schmitt & Schmitt, 2020). Repeated encounters with unknown lexical items in written input are crucial for successful learning to occur. Empirical evidence showing the importance of repetition in incidental L2 vocabulary learning from reading abounds (e.g., Elgort & Warren, 2014; Horst et al., 1998; Pellicer-Sánchez & Schmitt,

© The Author(s), 2022. Published by Cambridge University Press. This is an Open Access article, distributed under the terms of the Creative Commons Attribution licence (<http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted re-use, distribution and reproduction, provided the original article is properly cited.

2010; Uchihara *et al.*, 2019). Most studies suggest that considerable gains can occur after eight to ten encounters, but that different aspects of vocabulary knowledge (e.g., receptive and productive knowledge) might need a different number of exposures (Peters, 2020). However, the majority of studies exploring the role of frequency of exposure in vocabulary learning from reading have focused on the acquisition of single words. Despite the essential role of collocations (i.e., two or more words that co-occur more frequently than expected by chance alone), and other types of formulaic sequences, for achieving high levels of proficiency in a L2 (Siyanova-Chanturia & Pellicer-Sánchez, 2019), the examination of the incidental learning of formulaic sequences from reading and of the role of frequency of exposure in this process has only recently started to receive attention in the literature.

The evidence to date suggests that the form and meaning of collocations can be learnt incidentally from reading (Choi, 2017; Pellicer-Sánchez, 2017; Vilkaite, 2017; Webb & Chang, 2020) and reading-while-listening (Webb *et al.*, 2013), but the few studies exploring the role of frequency of exposure have yielded conflicting findings. Studies attest to a significant effect of repetition in the learning of collocations (e.g., Webb *et al.*, 2013), with some studies showing that just one extra exposure may lead to significant gains in collocational knowledge (e.g., Durrant & Schmitt, 2010). However, other studies have questioned this effect, indicating that extra exposures do not always lead to higher gains (e.g., Pellicer-Sánchez, 2017; Szudarski & Carter, 2016). Thus, our understanding of the role of frequency of exposure in the incidental learning of collocations from reading is still incomplete.

Crucially, recent eye-tracking research with single words has shown that through repeated encounters during reading, learners can develop familiarity with new forms and processing fluency, as reflected in a decrease in the number and durations of fixations (e.g., Elgort *et al.*, 2018; Godfroid *et al.*, 2018; Mohamed, 2018; Pellicer-Sánchez, 2016). However, no research to date has explored whether frequency of exposure has a similar effect on the processing of collocations. Eye-tracking research has shown that novel collocations initially attract more attention than known collocations (e.g., Choi, 2017), but processing changes throughout subsequent exposures are yet to be examined. While there is some evidence that suggests that small increases in frequency of exposure can lead to significantly higher gains in collocational knowledge (e.g., Durrant & Schmitt, 2010; Webb *et al.*, 2013), it is still unclear whether such frequency manipulations can lead to differences in the development of processing fluency. Finally, eye-tracking studies have recently suggested that processing time on lexical items during reading is related to their learning (e.g., Godfroid *et al.*, 2013; Mohamed, 2018; Pellicer-Sánchez, 2016); however, this relationship has only been examined with single words. Thus, the aim of the present investigation was to explore the role of frequency of exposure in the learning and processing of novel collocations over several encounters in a text, as well as the relationship between online processing of collocations and their learning.

Background

Collocations in a first and second language

Collocations, being one of different types of formulaic sequences (Wray, 2002), are often defined as a frequent co-occurrence of two or more words (Evert, 2008;

Sinclair, 1991). They are usually defined as the relation between a node and a collocate, “in which a collocate may be positioned either before or after the node and does not necessarily have to be adjacent to it” (Webb & Kagimoto, 2011, p. 262). They are syntagmatic units which can include both lexical and grammatical words (e.g., verb + noun [*pay tribute*], adjective + noun [*hot spice*], preposition + noun [*on guard*], and adjective + preposition [*immune to*], Henriksen, 2013, p. 29). Research has shown that a high percentage of language is made up of collocations. According to Hill (2000), “Collocations are found in up to 70% of everything we say, hear, read, or write” (p.53). Therefore, mastery of collocational competence is considered a crucial aspect of communicative competence, for both receptive and productive language competence (Cowie, 1992; Henriksen, 2013; Siyanova-Chanturia & Pellicer-Sánchez, 2019). The knowledge of collocations allows first language (L1) speakers to process language fluently (Siyanova-Chanturia & van Lancker, 2019) and to fulfill different communicative and pragmatic needs (Bardovi-Harlig, 2019; Kecskes, 2019). Thus, the knowledge of collocations is essential for L2 learners to reach high levels of proficiency (Bahns & Eldaw, 1993; Lewis, 2000) and to help learners communicate effectively (Hill, 2000). In addition, the knowledge of collocations is believed to foster both productive and receptive skills (Carter & McCarthy, 1988), including reading comprehension (Kremmel et al., 2017). Despite a central role of collocational knowledge in achieving fluent and successful communication in a L2, this linguistic phenomenon is acquired relatively slowly and is rarely mastered fully (Henriksen, 2013). L2 learners differ in their use of collocations compared to L1 speakers (Nguyen & Webb, 2017), both in terms of the type and amount of collocations used (Bahns & Eldaw, 1993; Granger, 1998; Laufer & Waldman, 2011; Lorenz, 1999; Serrano et al., 2015). This evidence has led researchers to claim that collocational competence is particularly difficult for L2 learners. However, research has only partially supported this assumption (Schmitt, 2010), with some studies suggesting that learners’ use of collocations may be similar to that of L1 speakers (Siyanova & Schmitt, 2008, Study 1; Omidian et al., 2017) and that they can be learnt at a rate similar to that of other aspects of lexical mastery (Pellicer-Sánchez, 2017; Webb & Kagimoto, 2009).

Incidental learning of collocations from reading

Incidental learning has been defined as the learning of linguistic features that accrues as a by-product of language usage without explicit intention to learn (Schmitt, 2010). Incidental learning is usually contrasted with intentional learning, when task instructions specifically ask learners to pay attention to the target items, and/or when learners are forewarned about the existence of vocabulary tests (Nation & Webb, 2011).

Reading has been shown to be an important source of incidental vocabulary learning in a L1 (Freebody & Anderson, 1983; Nagy et al., 1985; Saragi et al., 1978) and L2 (Pellicer-Sánchez & Schmitt, 2010; Waring & Takaki, 2003; Webb, 2007; Zahar et al., 2001). Although vocabulary research has traditionally focused on the acquisition of single words, recent years have seen an increase in the number of studies examining the acquisition of lexical items above the single word. Previous studies have shown that collocations can be intentionally learnt from different types

of activities (Boers *et al.*, 2014; Laufer, 2011; Laufer & Girsai, 2008; Peters, 2014; Webb & Kagimoto, 2009, 2011; see Pellicer-Sánchez & Boers, 2019, for a review). However, the number of studies exploring the incidental acquisition of collocations is still scarce.

Webb *et al.* (2013) showed that reading-while-listening to a modified graded reader led to a significant increase in learners' receptive knowledge of the form of the collocations. Sonbul and Schmitt (2013) extended these results to a reading-only context. They compared the effects of three learning conditions on the acquisition of adjective-noun collocations: 1) reading-only; 2) enhanced condition (reading a text with the target collocations enhanced); 3) and the decontextualized condition (target collocations embedded in activities). Results of the form recall and recognition post-tests showed that the reading-only condition led to a significant improvement in learners' collocational knowledge (27% for recall gains and 21% for recognition gains), with a significant advantage of the enhanced condition over the reading-only condition in both recall and recognition tests (37% and 32%, respectively). Similarly, Pellicer-Sánchez (2017) examined learning of collocations from repeated exposure in a text. L2 learners in this study read one of two versions of a story containing six adjective-pseudoword collocations that appeared either four or eight times. One week after the reading, learners' knowledge of the form, meaning, and collocation of the pseudowords were assessed. Results showed that reading led to a significant increase in learners' collocational knowledge, with learners in both groups being able to recognize 40.5% of the target collocations and to recall around 10% of them. The positive effect of reading on the acquisition of collocations has also been reported for non-adjacent collocations. The presence of intervening words between the components of the collocations does not seem to affect learning, with both adjacent and non-adjacent collocations being learned in a similar way (Vilkaite, 2017). Recent research has also shown that, when compared to other input modes, reading leads to similar gains to listening but to smaller gains than reading-while-listening (Webb & Chang, 2020).

Less optimistic results were reported by Szudarski (2012). L2 learners were exposed to a set of target collocations repeated six times each in a reading-only condition or a reading plus condition (reading plus activities on collocational patterns). Results showed that the improvement in collocational knowledge experienced by the reading-only group was not significantly different from that experienced by the control group. Similarly, Szudarski and Carter (2016) compared the effects of reading-only and enhanced reading (*i.e.*, reading condition with target words underlined) on the acquisition of verb-noun and adjective-noun collocations. Results showed that the reading-only condition did not improve learners' collocational knowledge. Researchers have claimed that the often-low intake of collocations, and other formulaic sequences, from reading is learners' failure to notice them in the input (Boers, 2020). As Boers (2020) explains, L2 readers might not pay attention to formulaic sequences during reading if they consist of familiar words (*e.g.*, *do your homework*) and when their meaning is transparent. Learners might attend to the part of the collocation that carries the most meaning, and the collocational patterning might be unnoticed (Boers, 2020). In fact, empirical evidence suggests that L2 learners prioritize individual words over collocational patterning (Hoang & Boers, 2016). If learners mainly attend to the part of a collocation that

carries the most meaning, learners would be expected to pay little attention to the known component in the known-unknown collocation type investigated in the present study.

Among the different factors affecting the incidental acquisition of new lexical items from reading, frequency of exposure, in particular, has received attention from vocabulary researchers. It is generally accepted that multiple encounters with collocations are needed for learners to develop collocational knowledge (Henriksen, 2013; Hill et al., 2000). However, only a few studies to date have looked at the effect of frequency of exposure on the incidental acquisition of collocations, and the studies available have produced mixed results. Durrant and Schmitt (2010) examined L2 learners' acquisition of adjective-noun collocations under three different conditions (single exposure, repeated exposure in the same sentences, and repeated exposure in different sentences). These authors showed that both repetition conditions outperformed the single exposure condition, suggesting that just one extra exposure was enough to observe significantly higher gains. Webb et al. (2013) also found that repeated exposures (1, 5, 10, and 15) had a positive effect on the incidental learning of collocations from reading-while-listening, with 15 exposures needed for sizeable learning to occur. Their findings suggested that five extra exposures were also sufficient to lead to significantly higher gains (as reflected in the comparison between 5–10 and 10–15 exposures). In contrast, other studies reported that a higher number of exposures did not always lead to better results at various levels of collocational competence. Pellicer-Sánchez (2017) compared the acquisition of adjective-pseudoword collocations from four vs. eight exposures and reported no significant differences between the two conditions. Similarly, the study by Szudarski and Carter (2016) showed that encountering the target collocations more frequently (6 vs. 12 exposures) did not lead to higher gains.

Eye-tracking and incidental vocabulary learning from reading

Eye-tracking has been used to investigate the incidental acquisition of L2 vocabulary from reading. Recent studies have shown that, when L2 readers encounter unknown lexical items, they spend more time processing them than matched known words (Godfroid et al., 2013; Mohamed, 2018; Pellicer-Sánchez, 2016). Number of encounters has a significant role in the processing of unknown words, with fixations and fixation durations decreasing as the number of exposures increases (Elgort et al., 2018; Godfroid et al., 2018; Mohamed, 2018; Pellicer-Sánchez, 2016). The processing of unknown lexical items starts to approximate that of known, control items by the eighth encounter (Elgort et al., 2018; Pellicer-Sánchez, 2016). These studies have also suggested a relationship between reading times on unknown words and vocabulary gains, with longer fixation durations linked to larger vocabulary gains (Godfroid et al., 2013; Mohamed, 2018; Pellicer-Sánchez, 2016). Pellicer-Sánchez (2016) also showed that advanced L2 readers' processing behavior was very similar to that of L1 readers, with the only difference being that the effect of frequency of exposure happened earlier for L1 readers than for L2 readers.

The above studies focused on the incidental acquisition of single words. To our knowledge, only one study to date has used eye-tracking to investigate incidental learning of collocations from reading. Choi (2017) examined the effect of textual

enhancement on the processing and learning of lexical (e.g., *big meal*) and grammatical collocations (e.g., *a bowl of cereal*). Participants read a passage containing 14 target collocations under an enhanced (i.e., target collocations boldfaced) or unenhanced condition. Results showed longer total reading times and more fixations in the enhanced condition than in the unenhanced condition. Eye-movement data showed that participants spent more time and made more fixations on the target collocations that were unknown to them than on those collocations that they already knew, in line with previous studies with individual words (Godfroid *et al.*, 2013; Mohamed, 2018; Pellicer-Sánchez, 2016). Both conditions led to significant learning of target collocations, with the enhanced condition outperforming the unenhanced condition.

The present study

The review of the literature suggests a wealth of findings, but also a number of notable gaps. First, the mixed results of previous studies question the effectiveness of reading-only conditions for the incidental acquisition of L2 collocational knowledge (Szudarski, 2012; Szudarski & Carter, 2016) and the role of frequency of exposure in this process (Szudarski & Carter, 2016; Pellicer-Sánchez, 2017). Secondly, the majority of previous studies have used offline post-reading tasks. Thus, our understanding of the processing of the different components of novel collocations throughout repeated exposures in reading and of the effect that different amounts of exposure may have on the development of processing fluency is still limited. Moreover, the relationship between online reading behavior and collocation learning has not been examined yet. Finally, no research to date has explored how L2 processing patterns compare to those of L1 readers when encountering novel collocations in reading. The present investigation aimed to address these gaps by asking the following questions:

1. Does frequency of exposure (four vs. eight exposures) have an effect on the learning of new words and their collocates from L1 and L2 reading?
2. How are novel collocations read throughout repeated exposures in L1 and L2 reading?
3. Are there any differences between the processing of novel collocations encountered four times versus those encountered eight times in L1 and L2 reading?
4. Is there a relationship between the online processing of novel collocations and learning, as measured by offline tests?

In order to address these questions, L1 and L2 speakers were asked to read a text containing novel adjective-pseudoword collocations (e.g., *magic salp*) while their eye movements were recorded and to complete a battery of tests assessing their knowledge of the pseudowords (nodes) and their collocates. Eye movements to the components of the collocations throughout the repeated encounters in the text were examined. The effect of frequency of exposure was also explored by comparing the learning and processing of novel collocations after having encountered them

four versus eight times in reading. A maximum of eight exposures were used as most studies have shown that considerable gains are achieved after 8–10 exposures. This was also the maximum number of exposures deemed possible given the length and duration of the experiment. The minimum number of repetitions was four to ensure that some learning could still take place, as research has shown that one or two exposures might not lead to detectable gains. Participants were randomly assigned to one of three conditions (*Target8*: eight occurrences of target collocations, *Target4*: four occurrences of target collocations, and *Control*: eight occurrences of control collocations). The combination of offline data (vocabulary post-tests) and eye-movement data allowed us to examine not only whether novel adjective-pseudoword combinations could be learnt from reading but also how the attention paid to the novel collocations (and their components) changed with repeated encounters, and whether the potential improvement in the processing speed differed after having seen the items four versus eight times. The combination of methods also allowed us to examine, for the first time, whether reading times on novel collocations predict their learning.

Method

Participants

Fifty-two L1 speakers of English (47 females; age range 18–20, $M = 19$, $SD = 0.7$) and 60 L2 learners of English (45 females, age range 19–39, $M = 26.5$, $SD = 4.2$) took part in the study. L1 participants were full-time undergraduate students at a university in the UK and received course credit, while L2 participants were post-graduate students at the same institution and received a small compensation. The L2 speakers of English, who came from a range of L1 backgrounds¹, were advanced learners who had met the English language entry requirement to study in the above institution. On average, these participants had spent 27.43 months in an English-speaking country, including their time in the UK (range in months 2–132, $SD = 23.2$); descriptive statistics by group are presented in the Supplementary materials S1). Prior to the experiment, the L2 participants completed a self-rating English proficiency test (Reading: $M = 8.2$, $SD = 1.2$; Writing: $M = 7.6$, $SD = 1.3$; Listening: $M = 7.9$, $SD = 1.2$; Speaking: $M = 7.4$, $SD = 1.5$). Statistical analysis revealed that there were no significant differences between the three conditions (*Target8*, *Target4*, *Control*) in the proficiency self-ratings, age of first contact with English, number of years living in an English-speaking country, and number of years learning English ($p > .1$, inferential statistics are presented in Supplementary materials S2). The research was carried out fulfilling ethical requirements in accordance with the standard procedures at the university where the experiment took place.

Target items

Because the study involved L1 and proficient L2 speakers of English, it was decided that collocations containing pseudowords would best allow us to investigate incidental learning of novel vocabulary from reading without needing to control for previous knowledge. Specifically, the study focused on adjective-pseudoword (pseudounoun) collocations. We were interested in examining the processing and learning of collocations in which the node was unknown, and the adjective was known to the participants.

Table 1. The pseudowords, control words, collocates, and their characteristics

| Pseudo-word | Length | Syllables | Control word | Length | Syllables | Frequency (Frequency band) | Common collocate | Frequency (Frequency band) |
|-------------|--------|-----------|--------------|--------|-----------|----------------------------|------------------|----------------------------|
| glabe | 5 | 1 | house | 5 | 1 | 49424 (1K) ^a | wooden | 3376 (1K) |
| nuse | 4 | 1 | bowl | 4 | 1 | 2363 (2K) | small | 43097 (1K) |
| redaster | 8 | 3 | criminal | 8 | 3 | 4995 (1K) | common | 19867 (2K) |
| salp | 4 | 1 | ring | 4 | 1 | 6757 (1K) | magic | 3114 (2K) |
| trobe | 5 | 1 | noise | 5 | 1 | 4408 (1K) | loud | 2340 (1K) |
| blaunts | 7 | 1 | clothes | 7 | 1 | 6952 (1K) | dirty | 2663 (1K) |

^aFrequency band: 1K = first 1,000 most frequency words in the corpus; 2K = first 2,000 most frequent words in the corpus.

First, a provisional pool of 63 pseudowords (varying in length and number of syllables) was selected from the ARC Nonword Database (<http://www.cogsci.mq.edu.au/research/resources/nwdb/nwdb.html>). They were selected based on the following criteria: neighborhood size (min = 1, max = 5), number of body neighbors (min = 1, max = 5), and number of phonological neighbors (min = 1, max = 5). In addition, no pseudohomophones were used, and only orthographically existing onsets and orthographically existing bodies, and only legal bigrams were extracted. Second, the pool of pseudowords was piloted with L1 speakers of English. Fifteen L1 speakers of English were asked to rate 63 pseudowords intermixed with 12 low-frequency real words (e.g., *burlap*, *scab*, *churn*) on a six-point Likert scale according to how English-like, or real word-like, their form was (1 = very unlikely to be an English word, and 6 = very likely to be an English word). Based on the scores provided, we selected six pseudowords whose mean score was at least 4 (*trobe*, *redaster*, *salp*: $M = 4$; *blaunts*: $M = 4.5$; *glabe*, *nuse*: $M = 5$). These items were thus judged as plausible English words in terms of their form.

As a second step, ten L1 speakers of English (not involved in the above procedure) were asked to rate the six pseudowords intermixed with seven low-frequency words (e.g., *stasis*, *vagary*, *aegis*) and seven other pseudowords on a six-point Likert scale according to how likely the connection between the item and a given meaning was (1 = very unlikely that this is the meaning of the pseudoword, and 6 = very likely that this is the meaning of the pseudoword). All target pseudowords had a mean score below 3 (*glabe*: $M = 1.5$; *trobe*: $M = 2.8$; *redaster*, *blaunts*: $M = 2$; *nuse*: $M = 2.5$; *salp*: $M = 1.8$), suggesting that there were no apparent connections between the form of the pseudowords and the meaning we attached to them.

The six pseudowords were used to replace six high-frequency real nouns in a story, which were the items for the control condition. The real nouns and the pseudowords were of the same length (in characters and syllables). Finally, we searched for common collocates (adjectives) of the nouns replaced, ensuring that the adjectives were sufficiently common (Table 1), the resulting collocations were sufficiently frequent (frequency of 13 or above occurrences in the British National Corpus) and

Table 2. The resulting pseudo-collocations, control collocations, and their characteristics

| Pseudo-collocation | Length | Control collocation | Length | Frequency | MI | Forward association | Backward association |
|--------------------|--------|---------------------|--------|-----------|-----|---------------------|----------------------|
| wooden glabe | 11 | wooden house | 11 | 14 | 3.1 | .05 | 0 |
| small nuse | 9 | small bowl | 9 | 40 | 5.3 | 0 | 0 |
| common redaster | 14 | common criminal | 14 | 13 | 3.7 | 0 | 0 |
| magic salp | 9 | magic ring | 9 | 13 | 5.9 | 0 | .01 |
| loud trobe | 9 | loud noise | 9 | 42 | 8.7 | .22 | .11 |
| dirty blaunts | 12 | dirty clothes | 12 | 26 | 7.1 | .01 | .01 |

statistically significant (Mutual Information score of 3 or above)², and that the forward (adjective → noun) and backward association strength (noun → adjective) was comparable (the Edinburgh Associative Thesaurus database was used, <http://www.eat.rl.ac.uk/>). The resulting pseudo- and control collocations contained the same adjective and were identical in length (Table 2).

Reading materials

The target and control collocations were inserted in a story previously designed and piloted by Pellicer-Sánchez (2016). This story was originally created to examine the incidental learning of six nouns repeated eight times throughout the story. For the purpose of the present study, those nouns were replaced by the adjective-pseudoword collocations. We kept the original meanings of the replaced words but selected a new set of better-controlled pseudowords and added the adjectives. Three versions of the original story were created. Two versions contained eight and four occurrences of the six pseudo-collocations, respectively (henceforth, *Target8* and *Target4* conditions). The third, control version contained eight occurrences of the six control collocations (henceforth, *Control* condition). The target collocations and their exposures were spread evenly across the text in the three versions. Minor changes were made to the original story to make sure that the collocations fitted in the context. The three versions of the story were 2,265 (*Target4*) and 2,324 (*Target8* and *Control*) words in length. They were presented on a computer screen in black over white background. The text was presented in Courier New font, size 18, with double spacing, across 25 screens; the text presented on each screen was of a similar length (eight lines, between 76–104 words). Each screen contained a maximum of two (different) pseudo-collocations. No target item appeared at the beginning or at the end of a line or a sentence (Rayner, 1977).

All words in the three versions of the story were from the first 4,000 most frequent words in English, except for the six target pseudowords, and three lower-frequency words (*someday*, *downstairs*, and *forever*). This meant a 2.2% of unknown words in versions *Target8* and *Control* and a 1.2% in version *Target4*. Given the advanced level of proficiency of learners, it was assumed that they would be familiar with the first 4,000 most frequent words in English, ensuring a lexical

coverage of about 98%. A true-false comprehension test containing 12 statements assessed participants' general comprehension of the story. None of the target items appeared in the comprehension questions. (Text for Target8 and Target4 versions available at <http://www.iris-database.org> and in the Supplementary materials S3 and S4).

Post-reading tests

Five vocabulary tests were used in the present study to assess the knowledge of the form and meaning of the pseudowords and their collocates. The tests were taken from Pellicer-Sánchez (2017) and modified by replacing the target items with the ones developed for the present study. The first test assessed participants' ability to recognize the form of the pseudoword from a set of four options provided. The second test measured participants' ability to recall the meaning of the pseudowords and was conducted by means of personal interviews (see Procedure section). Participants' ability to recognize the meaning of the target pseudowords was assessed in a multiple-choice test that presented the correct meaning of the target item and three options. An "I don't know option" was included in the recognition tests to minimize guessing effects. The distractors were all from the same word class and were semantically related to the content of the story to make them plausible options. The fourth test assessed participants' ability to recall the correct collocate of the target pseudowords and was conducted by means of interviews (see Procedure section). The last test assessed participants' ability to recognize the correct collocate from a set of five options (the correct collocate, three distractors, and an "I don't know" option). The distractors had the same word class and were all semantically related to the content of the story, to make the discrimination only possible when learners could recognize the collocate (vocabulary recognition tests available at <http://www.iris-database.org> and in the Supplementary materials S5). Having separate tests assessing knowledge of the pseudowords and the collocates allowed us to examine whether with this type of known-unknown combinations, readers would learn only the form and meaning of the pseudowords or would also learn the form of collocates that appeared with those pseudowords.

Procedure

The participants were randomly assigned to one of the three conditions (Target8, Target4, Control). Participants' eye movements (monocular recording, left eye) were monitored using a desktop-mounted, EyeLink 1000 Plus eye-tracker (SR Research, Canada), which samples data at 1000 Hz, with an accuracy of 0.25–0.5°, and a precision < 0.01°. Participants' head movements were minimized using a head rest. The experiment took place individually in a psycholinguistic laboratory. At the beginning of the session, L2 participants completed a language background questionnaire. The task was explained to the participants, and they were asked to read the story for comprehension in their most natural way. They were made aware of the comprehension questions that followed but were not informed about the presence of pseudowords in the story, or about the five vocabulary tests. The experiment started with a practice session (a five-page text and four True/False

comprehension questions) to familiarize participants with the procedure. A nine-point calibration took place before and after the practice session and halfway through the experiment. Drift correction was performed before each screen.

Participants pressed the “Enter” key to proceed from one screen to the next. They could not go back to previously read screens. After reading the story, participants responded to the True/False comprehension questions by pressing the “yes” or “no” buttons on the keyboard. Participants then completed the five vocabulary tests in the following order: pseudoword form recognition, pseudoword meaning recall, pseudoword meaning recognition, collocate recall, and collocate recognition. The form recognition, meaning recognition, and collocate recognition tests were completed individually on paper. The meaning recall and collocate recall tests were conducted by means of an interview. The researcher showed the participants cards with one pseudoword at a time and invited them to say everything they could remember about their meaning (in the meaning recall test) and about the adjectives or any other words that frequently occurred with the pseudowords (in the collocate recall test). The experiment took about 45 minutes to complete in conditions Target8 and Target4, and about 30 minutes in Control condition, as the latter did not involve the vocabulary post-tests.

Analysis

A score of 0 was given for incorrect responses and 1 for correct responses in the five vocabulary tests. Results of the two recall tests (pseudoword meaning recall and collocate recall) were scored following a strict procedure which did not give credit for partial knowledge. This was deemed important considering the small number of items. Only those instances where the expected response was provided were scored as correct. For example, if the word *pottery* was provided for the pseudoword *nuse* (“bowl”), it was considered incorrect, even though participants had shown a certain degree of semantic knowledge. The analysis of the offline measures included data from 31 L1 speakers (Target8 = 16, Target4 = 15) and 40 L2 speakers (Target8 = 20, Target4 = 20) (as participants in the control condition did not complete the vocabulary tests).

Regarding the eye-movement data, data from 41 L1 participants (Target8 = 12, Target4 = 14, Control = 15) and 45 L2 participants (Target8 = 15, Target4 = 15, Control = 15) were included in the analysis³. In addition, fixations shorter than 70 ms and longer than 1000 ms in any of the target and control adjectives and pseudowords were discarded. The data cleaning procedure resulted in 3.18% (or 115 data points) of L1 data, and 4.37% (or 228 data points) of L2 data being discarded. The lost data were equally distributed across the three conditions ($p > .1$). This produced a total of 3432 data points; 696 in the Target4 condition, 1296 in the Target8 condition, and 1440 in the Control condition.

The six adjective-pseudoword collocations in the Target8 and Target4 conditions and the six adjective-noun collocations in the Control condition were set as interest areas for the analysis. Following the approach suggested by Carrol and Conklin (2015), analyses were conducted with both the whole adjective-pseudoword combination as interest areas and with each component of the combination as a separate interest area. Examining the processing of each component allowed us to find out

whether any potential decrease in reading times and number of fixations with subsequent exposures of a collocation was driven by either component (the noun/pseudoword or the adjective collocate) or generally distributed over both. The eye-movement measures of interest were as follows: First Fixation Duration (the duration in milliseconds of the first fixation falling on an interest area); First-Run Reading Time (the sum of all fixation durations in milliseconds the first time an interest area was fixated, before exiting the region. This measure is often also called first-pass reading time, when it refers to a string longer than one word, and gaze duration, when it refers to one word in text); Total Reading Time (the sum of all fixation durations in milliseconds on the interest areas); and Fixation Count (the number of fixations on an interest area; Conklin *et al.*, 2018; Siyanova-Chanturia, 2013). First, Fixation Duration and First-Run Reading Time were chosen as early eye-movement measures that reflect lexical decoding processes. They provide information about learners' efforts to recognize and retrieve words from their lexicons. Total Reading Time and Fixation Count were chosen as aggregate late eye-movement measures that reflect word-to-text integration. As observed in previous vocabulary learning studies (e.g., Elgort *et al.*, 2018; Pellicer-Sánchez, 2016), a decrease in both early and late measures was expected as a function of exposure.

The eye-tracking data were analyzed via linear mixed-effect model run in R v3.6.1 (R Core Team, 2019), using the lme4 package v1.1-21 (Bates *et al.*, 2015). The random structure for models fitted to the whole dataset included random intercepts and random slopes for participants within exposure; the random slopes were dropped to simplify the structure and ensure convergence when subsetting the data by word class. Data from the vocabulary tests were analyzed via binary logistic regressions using the base-R stats package. Post hoc comparisons were performed with the emmeans package v1.6.0 (Lenth, 2021).

Results

Both L1 and L2 participants showed adequate comprehension of the story (L1 speakers: $M = 11$, $SD = 1.15$; L2 speakers: $M = 11.1$; $SD = 0.99$, $Max = 12$). There were no significant differences in the comprehension scores between L1 and L2 participants. Below, we first report on the analysis and results of the vocabulary test data and then the eye-tracking data.

Vocabulary tests

We first investigated whether significant differences existed between participant groups (L1 vs. L2) and between reading conditions (Target8 vs. Target4) in terms of participants' scores in the five vocabulary tests (pseudoword form recognition, pseudoword meaning recall, pseudoword meaning recognition, collocate recall, and collocate recognition). Table 3 presents the descriptive statistics for each of the five vocabulary measures.

A binary logistic regression was fitted to each vocabulary measure to test for the main effects of and interactions between participant group and condition. Table 4 shows significant main effects of participant group for all measures pertaining to the

Table 3. Post-reading vocabulary test scores for L1 and L2 speakers across the two frequency conditions (Target8 and Target4)

| Vocabulary test | L1 speakers <i>M</i> ^a (<i>SD</i>) | | L2 speakers <i>M</i> (<i>SD</i>) | |
|--------------------------------|---|-------------|------------------------------------|--------------|
| | Target8 | Target4 | Target8 | Target4 |
| Pseudoword form recognition | 5.53 (.74) | 5.53 (.74) | 4.90 (1.21) | 4.40 (1.27) |
| Pseudoword meaning recall | 2.80 (1.66) | 2.20 (1.74) | 1.85 (1.66) | 1.55 (1.191) |
| Pseudoword meaning recognition | 4.60 (1.30) | 4.20 (1.21) | 3.60 (1.47) | 3.55 (1.40) |
| Collocate recall | 2.47 (2.03) | 2.20 (1.42) | 1.90 (1.33) | 1.40 (1.31) |
| Collocate recognition | 4.60 (1.06) | 4.40 (0.99) | 4.05 (1.00) | 4.15 (1.27) |

^aMax score = 6

pseudonouns, but not the collocates, with scores of L2 speakers being significantly lower than those of L1 speakers. Importantly, reading condition did not appear to directly affect learning outcomes.

Eye-movement data

Analyses of the eye-tracking data were conducted first on the collocation (adjective + pseudoword) as a single interest area. Table 5 presents descriptive statistics for total reading time on the target and control collocations (Descriptive statistics for the remaining measures are included in the Supplementary materials S6).

The Total and First-Run Reading Time measures, as well as the First Fixation Duration measure, displayed the typical gamma distribution generally observed in eye-tracking (and, more generally, response-time) data. For this reason, these were log-transformed, achieving good residual normalization (one of the assumptions of linear mixed models). An initial maximal model was fitted with main effects of and interactions between participant group (L1 speakers vs. L2 speakers), condition (Target8, Target4, Control – with Control condition dummy-coded as baseline), and exposures (1-8 for Target8 and Control conditions, 1-4 for Target4 condition, with exposure 1 dummy-coded as baseline). The random structure included random intercepts and random slopes for exposure within participant, to account for potential between-subject differences in the rate of the predicted decrease in reading time with subsequent presentations of a collocation. Type III ANOVA with Satterthwaite approximation (Satterthwaite, 1946) was then run on the models to obtain omnibus statistics for the effects of the factors of interest. These results are presented in Table 6.

As expected, the results in Table 6 showed a main effect of participant group for Total Reading Time, First-Run Reading Times, and First Fixation Durations, with L2 speakers having longer reading times and fixation durations. These measures were also found to decrease as a function of exposure. A significant main effect of condition was found for all measures except First Fixation Duration, showing slower reading in conditions Target8 and Target4 relative to the control condition across exposures. Importantly, a significant interaction between condition and

Table 4. Summary of the binary logistic regressions to the 0–1 learning outcome measures using participant group (L1 vs. L2 speakers) and frequency condition (Target8 vs. Target4) as predictors

| Pseudoword form recognition | | | | |
|---|-------|------|-------|-------|
| Predictor | Est. | SE | z | p |
| ConditionTarget4 | 0.16 | 0.65 | 0.25 | .80 |
| ParticipantGroupL2_Speakers | −1.27 | 0.53 | −2.40 | .01* |
| ConditionTarget4: ParticipantGroupL2_Speakers | −0.35 | 0.74 | −0.47 | .63 |
| Pseudoword meaning recall | | | | |
| Predictor | Est. | SE | z | p |
| ConditionTarget4 | −0.37 | 0.32 | −1.16 | .24 |
| ParticipantGroupL2_Speakers | −0.79 | 0.32 | −2.40 | .01* |
| ConditionTarget4:ParticipantGroupL2_Speakers | 0.09 | 0.46 | 0.21 | .83 |
| Pseudoword meaning recognition | | | | |
| Predictor | Est. | SE | z | p |
| ConditionTarget4 | −0.33 | 0.37 | −0.90 | .36 |
| ParticipantGroupL2_Speakers | −0.93 | 0.35 | −2.64 | .008* |
| ConditionTarget4: ParticipantGroupL2_Speakers | 0.38 | 0.47 | 0.79 | .42 |
| Collocate recall | | | | |
| Predictor | Est. | SE | z | p |
| ConditionTarget4 | −0.09 | 0.32 | −0.27 | .78 |
| ParticipantGroupL2_Speakers | −0.45 | 0.33 | −1.36 | .17 |
| ConditionTarget4: ParticipantGroupL2_Speakers | −0.31 | 0.47 | −0.66 | .50 |
| Collocate recognition | | | | |
| Predictor | Est. | SE | z | p |
| ConditionTarget4 | −0.12 | 0.36 | −0.33 | .73 |
| ParticipantGroupL2_Speakers | −0.30 | 0.35 | −0.85 | .39 |
| ConditionTarget4: ParticipantGroupL2_Speakers | 0.12 | 0.48 | 0.25 | .80 |

exposure was also found for Total Reading Time and Fixation Count. Post hoc comparisons with Holm correction for multiple comparisons were performed on the interaction between condition and exposure. Given a large number of comparisons, we focus here on the comparisons at the initial and final exposures, that is, between the three conditions at the first and fourth exposures, between conditions Target8 and Control at the last (eighth) exposure, and between the last exposures in each condition (i.e., fourth in Target4 vs. eighth in Target8 vs. eighth in Control). These findings indicated a rapid habituation to the collocations in both Target8 and Target4 conditions (see Figure 1). As expected, at the first encounter, participants read control collocations faster than the pseudo-collocations in Target8 (Total Reading Time: $\beta = -0.61$, $SE = 0.07$, $z = -7.98$,

Table 5. Total reading time – Means (in ms) for the whole collocation for all conditions and each of the exposures (SD in brackets)

| | Exp 1 | Exp 2 | Exp 3 | Exp 4 | Exp 5 | Exp 6 | Exp 7 | Exp 8 |
|----------------|---------------|--------------|--------------|--------------|--------------|--------------|--------------|--------------|
| Target8 | | | | | | | | |
| L1 speakers | 780 (429) | 573 (329) | 490 (306) | 406 (215) | 431 (213) | 394 (181) | 369 (197) | 350 (170) |
| L2 speakers | 1041 (577) | 851 (464) | 701 (437) | 734 (446) | 623 (410) | 581 (291) | 596 (328) | 474 (225) |
| Target4 | | | | | | | | |
| L1 speakers | 700 (366) | 532 (297) | 485 (320) | 413 (237) | | | | |
| L2 speakers | 986 (564) | 862 (496) | 712 (406) | 556 (274) | | | | |
| Control | | | | | | | | |
| L1 speakers | 421 (207) | 412 (196) | 355 (165) | 372 (167) | 317 (191) | 332 (159) | 303 (193) | 298 (115) |
| L2 speakers | 568 (336) | 491 (238) | 463 (195) | 506 (372) | 450 (312) | 398 (182) | 368 (213) | 361 (175) |

$p = < .001$; Fixation Count: $\beta = -1.64$, $SE = 0.27$, $z = -5.92$, $p = < .001$) and Target4 (Total Reading Time: $\beta = -0.52$, $SE = 0.07$, $z = -6.96$, $p = < .001$; Fixation Count: $\beta = -1.17$, $SE = 0.27$, $z = -4.33$, $p = .003$) conditions. None of the comparisons at the fourth exposure were significant. Comparisons of the final presentation of the collocations in the three conditions showed that most of the comparisons were not significant either, suggesting that by the last exposure (eighth in conditions Target8 and Control, and fourth in condition Target4), novel adjective-pseudoword collocations were read in a similar way to control combinations, regardless of the condition. The comparisons at the final exposure also showed that target collocations were read in a similar way after having seen them four or eight times. The only difference was in Total Reading Time. The Total Reading Time for the collocations at the fourth exposure in condition Target4 was longer than the last (eighth) exposure of the control condition ($\beta = 0.33$, $SE = 0.07$, $z = 4.41$, $p = .002$). Total Reading Time also yielded significant differences between conditions Target8 and Control at exposures 5, 6, and 7, only converging at exposure 8 (Figure 1).

The subsequent analyses aimed at exploring the processing of each component of the collocations. The eye-tracking measures were recomputed for the two words in each collocation as distinct interest areas. Given the similar patterns observed for Total Reading Time and Fixation Count (Figure 1), for this second set of analyses the models were only fitted to Total Reading Time and First-Run Reading time as a late and early measure of reading, respectively.

An initial model was fitted to the dependent variables with participant group (L1 speakers vs. L2 speakers), condition (Target8, Target4, Control – with Control condition dummy-coded as baseline), word class (Noun vs. Adjective), and exposure (1-8 for conditions Target8 and Control, 1-4 for condition

Table 6. F and chi-square statistics for the effects and interactions from the maximal model fitted to Total Reading Time, First-Run Reading time, First Fixation Duration, and Fixation Count for the whole collocation. Effect sizes (η^2p and Cramer) and their 90% confidence intervals are also provided for significant effects and interactions

| Predictor | Total RT (log) | | First-Run RT (log) | | First Fixation (log) | | Fixation Count | |
|---------------------------------------|--------------------------------------|---------|--------------------------------------|---------|-------------------------------------|-------|--|---------|
| | F, (η^2p), [90% CI] | p | F, (η^2p), [90% CI] | p | F, (η^2p), [90% CI] | p | F, (v), [90% CI] | p |
| Participant Group | F(1,79) = 50.88, (0.39), [0.26,0.51] | <.0001* | F(1,76) = 48.81 (0.39), [0.25,0.50] | <.0001* | F(1,77) = 11.67 (0.13), [0.04,0.25] | .001* | $\chi^2(1,83) = 2.31$ | .12 |
| Condition | F(2,81) = 21.44 (0.35), [0.21,0.46] | <.0001* | F(2,79) = 3.79 (0.09), [0.01,0.19] | .02* | F(2,80) = 0.53 | .58 | $\chi^2(2,83) = 16.66 (0.38), [0.21,0.53]$ | <.0001* |
| Exposure | F(7,180) = 53.00 (0.67), [0.61,0.72] | <.0001* | F(7,179) = 14.62 (0.30), [0.21,0.36] | <.0001* | F(7,143) = 2.37 (0.10), [0.01,0.15] | .02* | $\chi^2(7,176) = 8.04$ | .32 |
| ParticipantGroup: Condition | F(2,81) = 1.36 | .26 | F(2,79) = 1.72 | .18 | F(2,80) = 0.48 | .61 | $\chi^2(2,83) = 0.09$ | .95 |
| ParticipantGroup: Exposure | F(7,180) = 0.73 | .64 | F(7,179) = 0.51 | .82 | F(7,143) = 0.36 | .92 | $\chi^2(7,176) = 2.69$ | .91 |
| Condition: Exposure | F(10,215) = 5.35 (0.20), [0.10,0.25] | <.0001* | F(10,150) = 1.08 | .37 | F(10,144) = 1.27 | .25 | $\chi^2(10,171) = 19.58 (0.30), [0.01,0.31]$ | .03* |
| ParticipantGroup: Condition: Exposure | F(10,215) = 1.20 | .29 | F(10,150) = 0.82 | .60 | F(10,144) = 1.82 | .06 | $\chi^2(10,171) = 10.69$ | .38 |

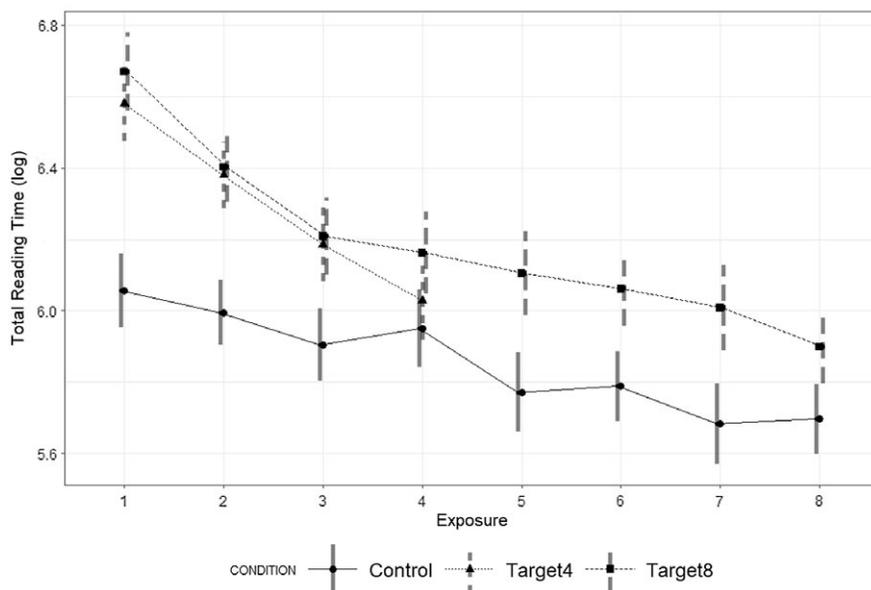


Figure 1. Interaction Between Condition (Target8, Target4, vs. Control) and Exposure (1–8) for the Total Reading Time (log) Measure for the Whole Collocation. Bars Indicate 95% Confidence Intervals.

Target4, with exposure 1 dummy-coded as baseline) as fixed effects with all interactions modeled. The random structure included random intercepts for the specific target words (e.g., “redaster” or “wooden”), to allow for the possibility that different words may be read differently at the first presentation, as well as random slopes for the effect of exposure within subject, to account for the possibility that Total Reading Time might change differently across subsequent presentations between participants. This model revealed a number of significant effects and interactions (presented in Table 7), including a three-way interaction between condition, word class, and exposure for both Total Reading Time and First-Run Reading Time.

In order to more easily interpret the observed three-way interactions and remove word class as a factor, the dataset was split by word class. However, given the smaller size of the resulting datasets, a model with a simpler random structure (with only random intercepts for each target word and for each participant) and excluding the fixed effect of word was fitted to the data. Table 8 reports the results of these models for both dependent variables and word classes. For the adjectives (Table 8), the main effect of condition revealed that total reading times, averaged across exposures and participant groups, were significantly longer in the target conditions than in the control condition (for both Target8, $t(1927) = 8.45$, $p < .0001$, and Target4, $t(881) = 5.17$, $p < .0001$), but that they did not differ between Target8 and Target4, $t(1236) = 1.45$, $p = 0.1$. The most notable effect was a significant interaction between participant group and exposure for the First-Run Reading Time measure, showing a steeper decrease across subsequent presentations but never catching up to the L1 speakers, who were already significantly faster than L2 speakers at baseline (Figure 2). The interaction between condition and exposure was not

Table 7. F statistics for the effects and interactions from the model fitted to Total Reading Time and First-Run Reading Time including word class as a fixed factor. Effect sizes (η^2p) and their 90% confidence intervals are also provided for significant effects and interactions

| Predictor | Total RT (log) | | First-Run RT (log) | |
|--|--------------------------------------|---------|--------------------------------------|---------|
| | F, (η^2p), [90% CI] | p | F, (η^2p), [90% CI] | P |
| Participant Group | F(1,69)=47.54 (0.38), [0.25,0.50] | <.0001* | F(1,60)=47.61 (0.44), [0.29,0.56] | <.0001* |
| Condition | F(2,83)=16.55 (0.42), [0.28,0.52] | <.0001* | F(2,73)=4.82 (0.17), [0.04,0.30] | .01* |
| Word Class | F(1,14)=5.38 (0.39), [0.03,0.65] | .03* | F(1,14)=6.07 (0.41), [0.05,0.66] | .02* |
| Exposure | F(7,147)=47.37 (0.75), [0.69,0.80] | <.0001* | F(7,240)=20.11 (0.40), [0.30,0.46] | <.0001* |
| ParticipantGroup:Condition | F(2,75)=1.97 | .14 | F(2,63)=2.36 | .10 |
| ParticipantGroup:Word Class | F(1,5473)=3.03 | .08 | F(1,5442)=5.47 | .01 |
| Condition:Word Class | F(2,41)=21.59 (0.02), [0.02,0.03] | <.0001* | F(2,56)=12.33 (0.02), [0.01,0.02] | <.0001* |
| ParticipantGroup:Exposure | F(7,146)=0.98 | .44 | F(7,240)=1.25 | .27 |
| Condition:Exposure | F(10,160)=4.37 (0.30), [0.13,0.37] | <.0001* | F(10,233)=1.17 | .31 |
| Word Class:Exposure | F(7,5453)=9.46 (0.01), [0.01,0.02] | <.0001* | F(7,5429)=7.84 (0.009), [0.00,0.01] | <.0001* |
| ParticipantGroup:Condition:Word Class | F(2,5466)=1.93 | .14 | F(2,5440)=2.39 | .09 |
| ParticipantGroup:Condition:Exposure | F(10,160)=1.84 (0.15), [0.01,0.19] | .05* | F(10,233)=0.92 | .51 |
| ParticipantGroup:Word Class:Exposure | F(7,5453)=0.97 | .44 | F(7,5429)=1.19 | .30 |
| Condition:Word Class:Exposure | F(10,5451)=3.80 (0.006), [0.00,0.01] | <.001* | F(10,5429)=2.62 (0.004), [0.00,0.01] | .003* |
| ParticipantGroup:Condition:Word Class:Exposure | F(10,5451)=1.41 | .16 | F(10,5429)=1.35 | .19 |

Table 8. F statistics for the effects and interactions from the model fitted to Total Reading Time and First-Run Reading Time for the adjective and the noun/pseudonoun data. Effect sizes (η^2p) and their 90% confidence intervals are also provided for significant effects and interactions

| Predictor | Adjective | | | | Noun/Pseudonoun | | | |
|-------------------------------------|----------------------------|---------|----------------------------|---------|---|---------|---|---------|
| | Total RT (log) | | First-Run RT (log) | | Total RT (log) | | First-Run RT (log) | |
| | F, (η^2p), [90% CI] | p | F, (η^2p), [90% CI] | p | F, (η^2p), [90% CI] | p | F, (η^2p), [90% CI] | p |
| ParticipantGroup | F(1,86)=30.86 | <.0001* | F(1,83)=28.16 | <.0001* | F(1,79)=29.85 (0.27), [0.14,0.40] | <.0001* | F(1,80)=35.96 (0.31), [0.18,0.43] | <.0001* |
| Condition | F(2,90)=7.37 | .0001* | F(2,86)=2.05 | .13* | F(2,42)=39.98 (0.49), [0.36,0.59] | <.0001* | F(2,60)=14.36 (0.26), [0.13,0.37] | <.001* |
| Exposure | F(7,2777)=17.06 | <.0001* | F(7,2773)=4.83 | <.0001* | F(7,2792)=60.74 (0.13), [0.11,0.15] | <.0001* | F(7,2793)=28.03 (0.07), [0.05,0.08] | <.0001* |
| ParticipantGroup:Condition | F(2,90)=0.66 | .51 | F(2,86)=0.65 | .52 | F(2,82)=1.77 | .15 | F(2,82)=2.30 | .10 |
| ParticipantGroup:Exposure | F(7,2777)=1.49 | .16 | F(7,2773)=2.02 | .04* | F(7,2792)=0.69 | .66 | F(7,2793)=0.73 | .64 |
| Condition:Exposure | F(10,2777)=0.93 | .50 | F(10,2773)=1.30 | .22 | F(10,2792)=9.22 (0.03), [0.02,0.04] | <.0001* | F(10,2793)=2.80 (0.009), [0.00,0.01] | .002* |
| ParticipantGroup:Condition:Exposure | F(10,2777)=1.13 | .33 | F(10,2773)=1.55 | .11 | F(10,2792)=2.24 (0.007), [0.00,0.01] | .009* | F(10,2793)=0.88 (0.003), [0.00,0.01] | .55 |

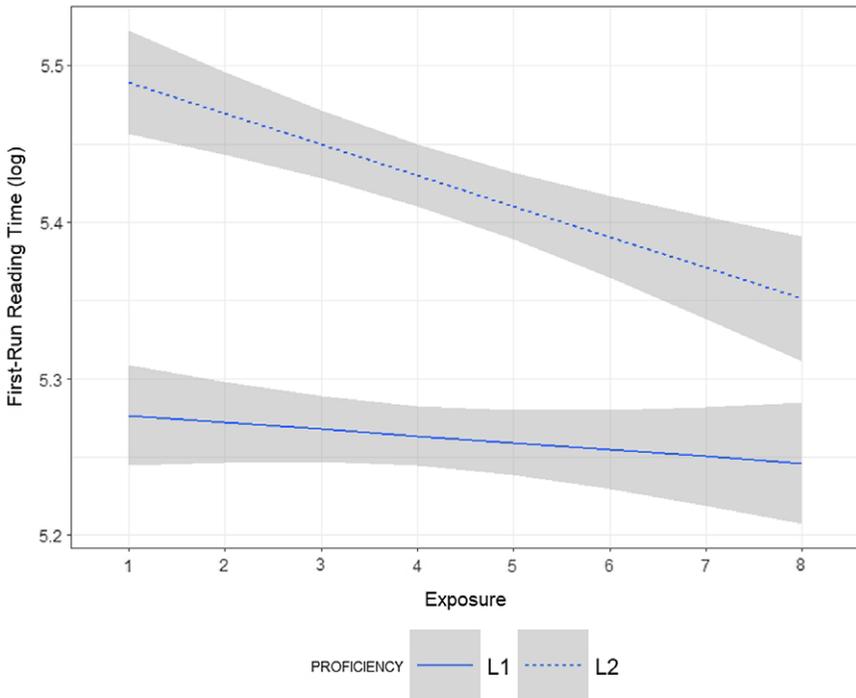


Figure 2. Interaction Between Participant Group (L1 vs. L2) and Exposure (1–8) for the First-Run Reading Time (log) Measure for the Adjectives. Shaded Areas Indicate 95% Confidence Intervals.

significant, indicating that the decrease in reading times across exposures did not differ by condition.

For the nouns/pseudonouns (Table 8), the pattern of results was more complex, with significant interactions between condition and exposure for both measures and a three-way interaction for Total Reading Time. Post hoc comparisons on this interaction showed that, at baseline (first exposure), reading times for control nouns (Control condition) were significantly shorter than for target pseudonouns for both participant groups in the Target8 (L1 speakers: $\beta = -0.77$, $SE = 0.12$, $z = -6.32$, $p < .0001$; L2 speakers: $\beta = -0.66$, $SE = 0.11$, $z = -5.77$, $p < .0001$) and Target4 (L1 speakers: $\beta = -0.68$, $SE = 0.11$, $z = -5.78$, $p < .0001$; L2 speakers: $\beta = -0.70$, $SE = 0.11$, $z = -6.06$, $p < .0001$) conditions, but that reading times on pseudonouns did not differ between Target4 and Target8 conditions. By the fourth exposure, reading times on the nouns did not differ between any of the conditions for L1 speakers, but for L2 speakers, reading times remained significantly different between Target8 and Target4 conditions ($\beta = -0.25$, $SE = 0.09$, $z = -2.69$, $p = .02$), as well as between Target8 and Control ($\beta = -0.54$, $SE = 0.11$, $z = -4.66$, $p < .0001$) and between Target4 and Control ($\beta = -0.28$, $SE = 0.11$, $z = -2.43$, $p = .04$) (see Figure 3). At the last exposure, L2 speakers spent longer time reading pseudonouns in Target8 condition than the control nouns ($\beta = -0.29$, $SE = 0.11$, $z = -2.50$, $p = .03$). None of the other comparisons were significant. First-Run Reading Time showed a similar pattern (Figure 4).

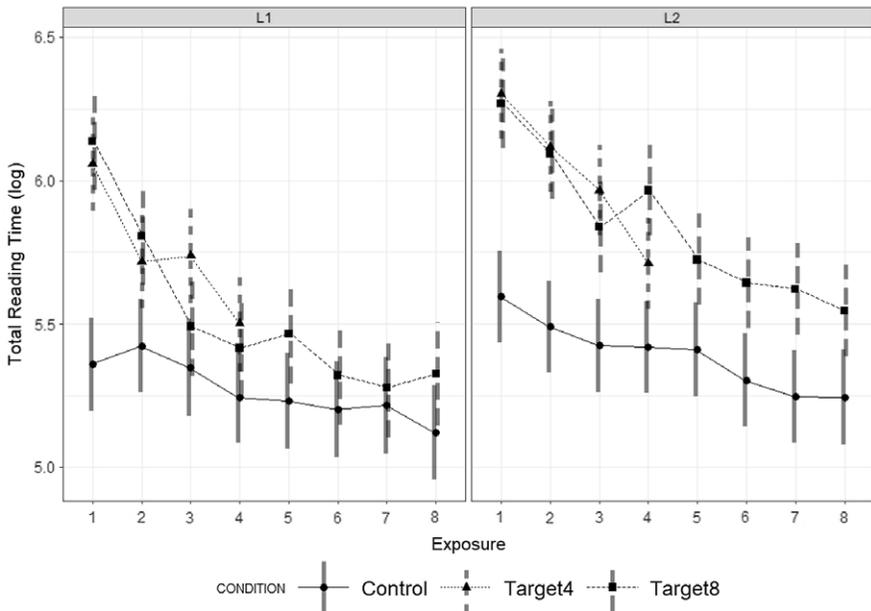


Figure 3. Interaction Between Participant Group, Exposure, and Condition (Control vs. Target8 vs. Target4) for Total Reading Time (log) for the Noun/Pseudonoun Components of Collocations. Bars Indicate 95% Confidence Intervals.

Lastly, we explored whether reading patterns across multiple exposures to the adjective and noun parts of the collocations could predict participants' learning accuracy for the form and meaning of the collocations. Because different outcomes relate to learning of different parts of the collocations (i.e., adjective collocate or pseudoword), the analyses were run by word class (i.e., noun vs. adjective). A cumulative reading time measure was computed by adding together the total reading times of all exposures to each target pseudoword or adjective and was used as a predictor, together with participant group, in a series of binary logistic regressions (since the data from individual trials were not used, and given the overall smaller number of data points, a regression was preferred to a linear mixed-effect model in this case). Log-transforming the reading time measure did not change the pattern of results, so the model presented here (Table 9) was fitted using the cumulative reading time measure in its natural scale (ms) for ease of interpretation. Across all other learning outcome variables, modeling an interaction between the predictors did not improve model fit, and only participant group was found to produce significant effects. Globally, these results show a lower learning performance for the L2 group compared to L1 speakers, but cumulative reading time did not significantly predict these differences.

General discussion

In the present investigation, we combined recordings of eye movements and a battery of post-reading vocabulary tests to examine the learning and processing of

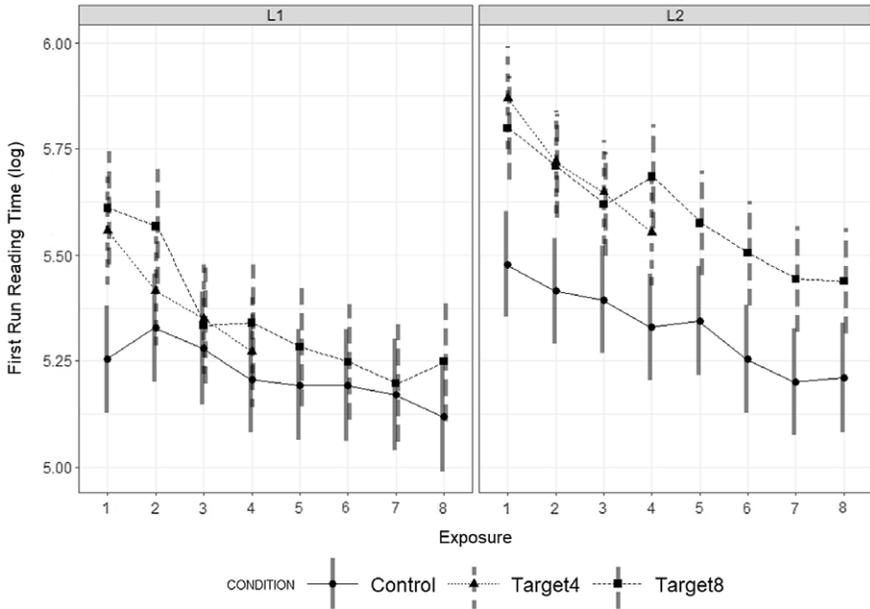


Figure 4. Interaction Between Participant Group, Exposure, and Condition (Control vs. Target8 vs. Target4) for First-Run Reading Time (log) for the Noun/Pseudonoun Components of Collocations. Bars Indicate 95% Confidence Intervals.

novel collocations in L1 and L2 reading, as well as the effect of frequency of occurrence on this process.

In response to the first research question (Does frequency of exposure have an effect on the learning of new words and their collocates in L1 and L2 reading?), our findings provided evidence that both L1 and L2 speakers can acquire knowledge of both new words and their collocates incidentally from repeated exposures in reading, in line with previous research (Pellicer-Sánchez, 2017; Sonbul & Schmitt, 2013). We found that not only did L1 and L2 readers acquire knowledge of the form and meaning of new words they encountered while reading, but they also gained knowledge of their recurring collocates. As expected, recognition scores were higher than recall scores, again, akin to previous studies on collocation learning (Pellicer-Sánchez, 2017; Peters, 2016; Szudarski, 2012; Webb et al., 2013). The results of this study have shown that, overall, L1 speakers could recall 39% of the correct collocates (Target8 = 41%; Target4 = 37%), while L2 learners were able to recall 27% (Target8 = 32%; Target4 = 23%). These recall figures are in line with those reported by Sonbul and Schmitt (2013) for advanced L2 learners (27%) and higher than those reported in Pellicer-Sánchez (2017) for intermediate L2 learners (11%). The recognition knowledge acquired by the L2 participants in this study is also higher than that reported in previous studies on incidental collocation learning, which ranged from 21 % (Sonbul & Schmitt, 2013) to around 50% (Pellicer-Sánchez, 2017; Webb et al., 2013). The comparison of the present results with those of Pellicer-Sánchez (2017) suggests a proficiency effect, with the more advanced learners in the current

Table 9. Summary of the binary logistic regression models fitted to the 0–1 learning outcome measures using cumulative reading time and participant group (L1 vs. L2) as predictors

| Pseudoword form recognition | | | | |
|--------------------------------|---------|--------|-------|--------|
| Predictor | Est. | SE | z | p |
| ParticipantGroupL2_Speakers | −1.51 | 0.38 | −3.90 | <.001* |
| Cumulative RT | 0.00006 | 0.0001 | 0.49 | .62 |
| Pseudoword meaning recall | | | | |
| Predictor | Est. | SE | z | p |
| ParticipantGroupL2_Speakers | −0.80 | 0.25 | −3.19 | .001* |
| Cumulative RT | 0.00009 | 0.0001 | 0.83 | .40 |
| Pseudoword meaning recognition | | | | |
| Predictor | Est. | SE | z | p |
| ParticipantGroupL2_Speakers | −0.69 | 0.25 | −2.74 | .005* |
| Cumulative RT | 0.00009 | 0.0001 | 0.83 | .40 |
| Collocate recall | | | | |
| Predictor | Est. | SE | z | p |
| ParticipantGroupL2_Speakers | −0.53 | 0.25 | −2.08 | .03* |
| Cumulative RT | −0.0001 | 0.0001 | −0.65 | .51 |
| Collocate recognition | | | | |
| Predictor | Est. | SE | z | p |
| ParticipantGroupL2_Speakers | −0.10 | 0.26 | −0.38 | .7 |
| Cumulative RT | −0.0002 | 0.0001 | −1.36 | .17 |

study achieving higher gains. The results of the offline measures have also shown that the scores are significantly lower for L2 speakers than for L1 speakers. Importantly, the lack of a main effect of condition on the vocabulary scores across both participant groups questions the effect of frequency of exposure on incidental learning of collocations. This is in line with the results of Pellicer-Sánchez (2017) and Szudarski and Carter (2016) but is contrary to other studies that have shown a significant effect of exposure (Durrant & Schmitt, 2010; Webb et al., 2013). This shows that increased exposure does not necessarily improve learning, at least when the difference between the number of exposures is relatively small (i.e., four vs. eight). The differences between the results of this study and those of earliest studies, both in terms of the benefits of reading and the effect of frequency of exposure, are likely due to the different nature of the target lexical items selected, the different assessment instruments used, the different number of exposures provided (ranging from 1 vs. 2 exposures in Durrant & Schmitt, 2010, to 15 exposures in Webb et al., 2013), and the length of the reading task. A longer reading text and a different number of target items might have revealed a significant effect of frequency.

In response to the second research question (How are novel collocations read throughout repeated exposures in L1 and L2 reading?), the analyses of the whole collocations showed that at the first encounter, both L1 and L2 speakers read novel collocations more slowly than controls, in line with previous studies with single words in L1 and L2 contexts (Chaffin *et al.*, 2001; Godfroid *et al.*, 2013; Pellicer-Sánchez, 2016) and with L2 collocation studies (Choi, 2017). After the first exposure, there was a decrease in the number and durations of fixations across subsequent exposures, also supporting the findings of previous studies with single words (Elgort *et al.*, 2018; Godfroid *et al.*, 2018; Mohamed, 2018; Pellicer-Sánchez, 2016). Differences in the processing of target and control collocations started to disappear by the fourth exposure in some of the measures analyzed. By the eighth exposure, target collocations were read in a similar way to control collocations.

Interesting differences emerged regarding the manner in which the collocation components were processed. Results showed that L1 and L2 speakers read both the pseudonoun and the adjective, but processing patterns differed. Concerning the processing of the node of the collocation – the pseudowords – results were similar to the patterns reported for the whole collocation. At the first encounter, readers took longer to process the pseudowords in Target8 and Target4 conditions than the control words. There was a significant decrease in the number and durations of fixations after this first exposure. By the fourth exposure, there were no differences across conditions for L1 speakers, while L2 speakers still took longer to process them than control words (in line with Pellicer-Sánchez, 2016). Interestingly, by the eighth encounter, L2 speakers still processed the pseudowords in Target8 condition more slowly than control words (and almost at significance level when comparing the last exposures in Target8 and Target4 conditions). This is interesting as it goes against previous research showing no difference between the processing of target and control words by the eighth encounter (Elgort *et al.*, 2018; Pellicer-Sánchez, 2016). The potential attempts to link the pseudowords with their collocates, as reflected in the collocate tests, might have interfered with the processing, slowing down the processing of the pseudoword. It might be the case that L2 readers might need more exposures to develop this speed and processing fluency when novel words are accompanied by a collocate.

Importantly, when examining the processing of the adjective collocate, results showed that readers spent more time processing the adjectives in Target8 and Target4 conditions than in the control condition. This suggests that readers paid attention to the collocate, not only to the unknown component, and that more attention was paid to the adjective when it appeared as part of an unknown collocation than when it was part of a known combination. This might reflect attempts to integrate the collocate to the pseudonouns in Target8 and Target4 conditions. Crucially, as suggested by the lack of interaction between condition and exposure, the decrease in reading times did not differ by condition. This analysis also showed that the component that was driving the significant interaction between condition and exposure in the analysis of the whole collocation was the pseudoword and provides further evidence for the need to look at individual collocation constituents as well as at the entire sequence in eye-tracking studies (Carrol & Conklin, 2015; Jiang *et al.*, 2020).

Interestingly, few differences were observed in the processing patterns of L1 and L2 speakers. As reported above, L2 speakers were generally slower when processing the whole collocations, and their reading times on target pseudonouns by the eighth exposure were still longer than those on control words. The lack of other differences between the two participant groups could be due to the advanced proficiency of the L2 group, as well as to the experimental materials used. A more challenging text and a larger number of target items might have yielded other significant differences.

In response to the third research question (Are there any differences between the processing of novel collocations encountered four times versus those encountered eight times in L1 and L2 reading?), results showed that target collocations (and the pseudowords) were processed in a similar way after being seen eight times (Target8 condition) and four times (Target4 condition), as revealed in the comparison of the last exposure in both conditions. This suggests that the frequency manipulation in the present study (four vs. eight) did not seem to make a difference in the development of speed and fluency with which the new items were read. This conclusion, based on the analysis of the online data, is further supported by the lack of frequency effect found in the analysis of the offline post-reading test data. This further supports recent claims in the literature that the number of encounters is not the only, or the major, factor known to affect incidental vocabulary learning from reading, as other factors such as relevance, salience, and distribution, also play a role (e.g., Laufer & Rozovski-Roitblat, 2015; Szudarski & Carter, 2016; Webb & Chang, 2015). In their meta-analysis, Uchihara et al. (2019) showed that the significant effect of frequency of exposure on incidental vocabulary learning was moderated by learner (age, vocabulary knowledge), treatment (spaced learning, visual support, engagement, range in number of encounters), and methodological (nonword use, forewarning of upcoming comprehension test, vocabulary test format) variables. The use of pseudowords as the node of a collocation, as well as the salience of the repeated items in a short treatment, is likely to have moderated the effect of frequency in the present study. It is important to note, though, that a different frequency manipulation (e.g., 2 vs. 8, or 4 vs. 10) might have yielded significant effects of frequency. Different frequency manipulations should thus be explored in future research.

Finally, in response to the fourth research question (Is there a relationship between the online processing of novel collocations and learning, as measured by offline tests?), the results of the present investigation revealed no significant relationship between the processing of collocations and learning gains for any of the participant groups. This finding goes against the results of previous studies with single words which indicated that increased attention to target items was positively related to learning gains (Godfroid et al., 2013; Godfroid et al., 2018; Mohamed, 2018; Pellicer-Sánchez, 2016). However, it is in line with results of the study by Elgort et al. (2018) which also failed to find a reliable relationship. This lack of connection between reading times on words and vocabulary scores also supports the findings of studies in other vocabulary learning contexts (e.g., Bisson et al., 2015). While several incidental vocabulary learning studies have reported a positive relationship between longer reading times and improved performance in vocabulary tests, other studies have shown that longer reading times were connected to lower vocabulary scores, with longer reading times representing processing difficulties

(Montero Perez *et al.*, 2015). Results of this study provide further evidence for the complexity of the relationship between eye movements and performance measures (Pellicer-Sánchez, 2020), and the possibility that the number and duration of fixations on novel lexical items may reflect different subprocesses (Godfroid *et al.*, 2018), calling for further research in the area.

While the results of this study are novel in that they address a notable gap in the literature, it is important to note that they are primarily applicable to one type of collocation – adjective-noun collocations where the adjective collocates is known and the noun is unknown. The online processing and learning of items of different characteristics (e.g., different parts of speech, both vs. one unknown component, idiomatic vs. literal meaning) may show a different pattern. Future research should thus explore the effect of exposure on the learning and processing of different types of word combinations. Notably, the use of pseudowords in place of high-frequency real words (as in the present study) might mean that, when encountering the pseudowords, participants accessed the meaning of the high-frequency word against their existing collocational knowledge. The knowledge demonstrated in the test might then be a reflection of the knowledge acquired from reading, as well as of the existing collocational knowledge of the high-frequency word. Exposure to the pseudowords in the first three tests could have also prompted knowledge of the collocates that was assessed in later tests. It is also important to note that the high-recognition scores reported in this study could be due to the small number of target items. Future studies should consider having a longer experimental text that would allow for the inclusion of a higher number of target items while maintaining the required lexical coverage. Furthermore, while the learning condition examined in the present study was considered incidental, the repeated exposure to the target items in a relatively short text may have rendered the target items rather salient for the readers, making the learning more intentional than incidental, thus influencing processing patterns and learning gains. The results reported could have also been affected by participant's characteristics that were not controlled for in the present study, such as gender and L1 background. Future studies should also use measures of proficiency that involve objective measurements, instead of the self-ratings used in the present study, and include these as variables in the analyses. It will also be interesting in future research to examine the role that important individual differences, such as working memory capacity, have on the processing of novel collocations.

Conclusions

The results of the present study have shown that repeated exposures in L1 and L2 reading led to improved knowledge of unknown words and their collocates, as well as to improved fluency with which the novel adjective-pseudoword combinations were read, with reading times decreasing as a function of exposure. These findings further attest to the benefits of repeated exposures in reading for the development of several aspects of vocabulary knowledge. The analysis of eye movements on the whole collocations suggests that eight exposures were enough to develop processing speed and to be able to read novel collocations in a way similar to controls, with L1 speakers developing processing fluency earlier than L2 learners. However, the

analysis of the unknown component of the collocation (i.e., the pseudoword) showed that for L2 speakers, eight exposures were not enough to process unknown pseudowords similar to known words, suggesting that more exposures are needed to develop processing fluency when unknown words are embedded in this type of known-unknown collocations. Importantly, the frequency manipulation in the present study (four vs. eight) did not seem to lead to major differences in the learning or processing of known-unknown collocations. Cumulative reading times were not a predictor of vocabulary gains, providing further evidence for the complex relationship between eye movements and outcome measures and calling for further research in this area.

Supplementary material. For supplementary material accompanying this paper visit <https://doi.org/10.1017/S014271642200011X>.

Acknowledgements. We are grateful to the editor and the three anonymous reviewers for their insightful comments on earlier versions of the manuscript.

Notes

1. Breakdown of L1 backgrounds: Spanish = 9; Chinese = 6; Arabic = 6; German = 6; Italian = 5; Greek = 5; Turkish = 4; French = 3; Malay = 3; Bengali = 2; Czech = 1; Divehi = 1; Finnish = 1; Indonesian = 1; Kurdish = 1; Polish = 1; Romanian = 1; Russian = 1; Thai = 1; Ukrainian = 1.
2. Mutual Information is a measure of the strength of association between two (or more) words.
3. Data of poor quality (i.e., data which contained sizeable track loss or problematic drift) were not included in the analysis. Data were initially inspected with the temporal graph and spatial overlay view in DataViewer to identify trials where track loss or problematic drift had occurred. Given the small number of trials per participant (25 trials), data from participants when sizeable track loss occurred in more than two trials were not included in the analysis. While this is a rather conservative approach, it allowed us to ensure data quality and reliability of results. This resulted in data from 41 L1 and 45 L2 participants being included in the analysis.

References

- Bahns, J. & Eldaw, M. (1993). Should we teach EFL students collocations? *System*, **21**, 101–114. Doi: [10.1016/0346-251X\(93\)90010-E](https://doi.org/10.1016/0346-251X(93)90010-E)
- Bardovi-Harlig, K. (2019). Formulaic language in second language pragmatics research. In A. Siyanova-Chanturia & A. Pellicer-Sánchez (Eds.), *Understanding formulaic language: A second language acquisition perspective* (pp. 97–114). New York: Routledge.
- Bates, D., Maechler, M., Bolker, B., & Walker, S. (2015). Fitting linear mixed-effects models using lme4. *Journal of Statistical Software*, **67**(1), 1–48. Doi: [10.18637/jss.v067.i01](https://doi.org/10.18637/jss.v067.i01)
- Bisson, M-J., Van Heuven, W., Conklin, K., & Tunney, R. (2015). The role of verbal and pictorial information in multi-modal incidental acquisition of foreign language vocabulary. *Quarterly Journal of Experimental Psychology*, **68**(7), 1306–1326. Doi: [10.1080/17470218.2014.979211](https://doi.org/10.1080/17470218.2014.979211)
- Boers, F. (2020). Factors affecting the learning of multiword items. In S. Webb (Ed.), *The Routledge Handbook of Vocabulary Studies* (pp. 143–157). New York: Routledge.
- Boers, F., Demecheleer, M., Coxhead, A., & Webb, S. (2014). Gauging the effects of exercises on verb-noun collocations. *Language Teaching Research*, **18**, 54–74. Doi: [10.1177/1362168813505389](https://doi.org/10.1177/1362168813505389)
- Carrol, G., & Conklin, K. (2015). Eye-tracking multi-word units: some methodological questions. *Journal of Eye Movement Research*, **7**(5), 1–11. Doi: [10.16910/jemr.7.5.5](https://doi.org/10.16910/jemr.7.5.5)
- Carter, R. & McCarthy, M. (1988). *Vocabulary and language teaching*. New York: Longman.
- Chaffin, R., Morris, R. K., & Seely, R. E. (2001). Learning new word meanings from context: A study of eye movements. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, **27**, 225–235. Doi: [10.1037/0278-7393.27.1.225](https://doi.org/10.1037/0278-7393.27.1.225)
- Choi, S. (2017). Processing and learning of enhanced English collocations: An eye movement study. *Language Teaching Research*, **21**(3), 403–426. Doi: [10.1177/1362168816653271](https://doi.org/10.1177/1362168816653271)

- Conklin, K., Pellicer-Sánchez, A., & Carrol, G.** (2018). *Eye-tracking: A guide for Applied Linguistics Research*. Cambridge: Cambridge University Press.
- Cowie, A.** (1992). Multiword lexical units and communicative language teaching. In P. Arnaud and H. Bejoint (Eds.), *Vocabulary and applied linguistics* (pp. 216–331). London: Macmillan Academic and Professional LTD.
- Durrant, P. & Schmitt, N.** (2010). Adult learners' retention of collocations from exposure. *Second Language Research*, *26*(2), 163–188. Doi: [10.1177/0267658309349431](https://doi.org/10.1177/0267658309349431)
- 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*, 341–366. Doi: [10.1017/S0272263117000109](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. Doi: [10.1111/lang.12052](https://doi.org/10.1111/lang.12052).
- Evert, S.** (2008). Corpora and collocations. In A. Ludeling & M. Kyto (Eds.), *Corpus linguistics: An international handbook* (pp. 1212–1248). Berlin: Mouton de Gruyter.
- Freebody, P., & Anderson, R.** (1983). Effects of vocabulary difficulty, text cohesion, and schema availability on reading comprehension. *Reading Research Quarterly*, *18*, 277–294.
- Godfroid, A., Ahn, J., Choi, I., Ballard, L., Cui, Y., Johnston, S., Lee, S., Sarkar, A. & Yoon, H.** (2018). Incidental vocabulary learning in a natural reading context: An eye-tracking study. *Bilingualism: Language and Cognition*, *21*(3), 563–584. Doi: [10.1017/S1366728917000219](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*, 483–517. Doi: [10.1017/S0272263113000119](https://doi.org/10.1017/S0272263113000119)
- Granger, S.** (1998). Prefabricated patterns in advanced EFL writing: Collocations and formulae. In A. P. Cowie (Ed.), *Phraseology: Theory, analysis and application* (pp. 145–160). Oxford: Oxford University Press.
- Henriksen, B.** (2013). Research on L2 learners' collocational competence and development. In C. Bardel, C. Lindquist & B. Laufer (Eds.), *L2 Vocabulary acquisition, knowledge, and use: New perspectives on assessment and corpus analysis*. *EuroSLA Monograph Series 2* (pp. 29–56). European Second Language Association.
- Hill, J.** (2000). Revising priorities: from grammatical failure to collocational success. In M. Lewis, (Ed.), *Teaching Collocation* (pp. 47–69). Hove: Language Teaching Publications.
- Hill, J., Lewis, M., & Lewis, M.** (2000). Classroom strategies, activities and exercises. In M. Lewis, (Ed.), *Teaching collocations: further developments in the lexical approach* (pp. 88–117). Boston, MA: Thomson.
- Hoang, H., & Boers, F.** (2016). Re-telling a story in a second language: How well do adult learners mine an input text for multiword expressions? *Studies in Second Language Learning and Teaching* *6*(3), 513–535. Doi: [10.14746/ssl.2016.6.3.7](https://doi.org/10.14746/ssl.2016.6.3.7)
- Horst, M., Cobb, T., & Meara, P.** (1998). Beyond a clockwork orange: Acquiring second language vocabulary through reading. *Reading in a Foreign Language*, *11*, 207–223.
- Jiang, S., Jiang, X., & Siyanova-Chanturia, A.** (2020, first view). Multi-word expression processing in children: An eye-tracking study of Chinese. *Applied Psycholinguistics*.
- Kecskes, I.** (2019). Formulaic language and its place in intercultural pragmatics. In A. Siyanova-Chanturia, & A. Pellicer-Sánchez (Eds.), *Understanding formulaic language: A second language acquisition perspective* (pp. 170–188). New York: Routledge.
- Kremmel, B., Brunfaut, T., & Alderson, C.** (2017). Exploring the role of phraseological knowledge in foreign language reading. *Applied Linguistics*, *38*(6), 848–870. Doi: [10.1093/applin/amv070](https://doi.org/10.1093/applin/amv070)
- Laufer, B.** (2011). The contribution of dictionary use to the production and retention of collocations in a second language. *International Journal of Lexicography*, *24*, 29–49.
- Laufer, B., & Girsai, N.** (2008). Form-focused instruction in L2 vocabulary learning. *Applied Linguistics*, *29*, 694–716. Doi: [10.1093/applin/amn018](https://doi.org/10.1093/applin/amn018)
- Laufer, B., & Rozovski-Roitblat, B.** (2015). Retention of new words: Quantity of encounters, quality of task, and degree of knowledge. *Language Teaching Research*, *19*, 687–711. Doi: [10.1177/1362168814559797](https://doi.org/10.1177/1362168814559797)
- Laufer, B., & Waldman, T.** (2011). Verb-noun collocations in second-language writing: A corpus analysis of learners' English. *Language Learning*, *61*(2), 647–672. Doi: [10.1111/j.1467-9922.2010.00621.x](https://doi.org/10.1111/j.1467-9922.2010.00621.x)
- Lenth, R. V.** (2021). Emmeans: Estimated Marginal Means, aka Least-Squares Means.

- Lewis, M. (Ed.). (2000). *Teaching collocations: Further developments in the lexical approach*. Boston, MA: Thomson.
- Lorenz, T. R. (1999). *Adjective intensification – learners versus native speakers: A corpus study of argumentative writing*. Amsterdam: Rodopi.
- Mohamed, A. A. (2018). Exposure frequency in L2 reading. An eye-movement perspective of incidental vocabulary learning. *Studies in Second Language Acquisition*, *40*, 269–293. Doi: [10.1017/S0272263117000092](https://doi.org/10.1017/S0272263117000092).
- Montero Perez, M., Peters, E., & Desmet, P. (2015). Enhancing vocabulary learning through captioned video: An eye-tracking study. *The Modern Language Journal*, *99*(2), 308–328. Doi: [10.1111/modl.12215](https://doi.org/10.1111/modl.12215).
- Nagy, W., Herman, P., & Anderson, R. (1985). Learning words from context. *Reading Research Quarterly*, *20*, 233–53.
- Nation, I.S.P., & Webb, S. (2011). *Researching and analyzing vocabulary*. Boston, MA: Heinle, Cengage Learning.
- Nguyen, T. M. H., & Webb, S. (2017). Examining second language receptive knowledge of collocation and factors that affect learning. *Language Teaching Research*, *21*(3), 298–320. Doi: [10.1177/1362168816639619](https://doi.org/10.1177/1362168816639619).
- Omidian, T., Beliaeva, N., Todd, L., and Siyanova-Chanturia, A. (2017). The use of academic words and formulae in L1 and L2 secondary school writing. *New Zealand Studies in Applied Linguistics*, *23*(2), 39–59.
- Pellicer-Sánchez, A. (2016). Incidental L2 vocabulary acquisition from and while reading: An eyetracking study. *Studies in Second Language Acquisition*, *38*, 97–130. Doi: [10.1017/S0272263115000224](https://doi.org/10.1017/S0272263115000224)
- Pellicer-Sánchez, A. (2017). Incidental learning of L2 collocations: A classroom study. *Language Teaching Research*, *21*(3), 381–402. Doi: [10.1177/1362168815618428](https://doi.org/10.1177/1362168815618428).
- Pellicer-Sánchez, A. (2020). Expanding English vocabulary knowledge through reading: Insights from eye-tracking studies. *RELC Journal*, *51*(1), 134–146. Doi: [10.1177/0033688220906904](https://doi.org/10.1177/0033688220906904)
- Pellicer-Sánchez, A., & Boers, F. (2019). Pedagogical approaches to the teaching and learning of formulaic language. In A. Siyanova-Chanturia & A. Pellicer-Sánchez (Eds.), *Understanding formulaic Language: A second language acquisition perspective* (pp. 153–173). New York: Routledge.
- Pellicer-Sánchez, A., & Schmitt, N. (2010). Incidental vocabulary acquisition from an authentic novel: Do things fall apart? *Reading in a Foreign Language*, *22*, 31–55.
- Peters, E. (2014). The effects of repetition and time of post-test administration on EFL learners' form recall of single words and collocations. *Language Teaching Research*, *18*, 75–94. Doi: [10.1177/1362168813505384](https://doi.org/10.1177/1362168813505384)
- Peters, E. (2016). The learning burden of collocations: The role of interlexical and intralexical factors. *Language Teaching Research*, *20*(1), 113–138. Doi: [10.1177/1362168814568131](https://doi.org/10.1177/1362168814568131)
- Peters, E. (2020). Factors affecting the learning of single words. In S. Webb (Ed.), *The Routledge Handbook of Vocabulary Studies* (pp. 125–142). New York: Routledge.
- Rayner, K. (1977). Visual attention in reading: Eye movements reflect cognitive processes. *Memory & Cognition*, *4*, 443–448.
- R Core Team (2014). *R: A language and environment for statistical computing*. R Foundation for Statistical Computing, Vienna, Austria. URL <http://www.R-project.org/>
- Saragi, T., Nation, P., & Meister, F. (1978). Vocabulary learning and reading. *System*, *6*, 72–78.
- Schmitt, N. (2010). *Researching vocabulary: A vocabulary research manual*. London: Palgrave Macmillan.
- Schmitt, N. & Schmitt, D. (2020). *Vocabulary in language teaching* (2nd Edn.). Cambridge: Cambridge University Press.
- Serrano, R., Stengers, H., & Housen, A. (2015). Acquisition of formulaic sequences in intensive and regular EFL programmes. *Language Teaching Research*, *19*, 89–106. Doi: [10.1177/1362168814541748](https://doi.org/10.1177/1362168814541748).
- Sinclair, J. (1991). *Corpus, concordance, collocation*. Oxford: Oxford University Press.
- Siyanova, A., & Schmitt, N. (2008). L2 learner production and processing of collocation: A multistudy perspective. *The Canadian Modern Language Review*, *64*, 429–458. Doi: [10.3138/cmlr.64.3.429](https://doi.org/10.3138/cmlr.64.3.429)
- Siyanova-Chanturia, A. (2013). Eye-tracking and ERPs in multi-word expression research: A state-of-the-art review of the method and findings. *The Mental Lexicon*, *8*(2), 245–268. Doi: [10.1075/ml.8.2.06siy](https://doi.org/10.1075/ml.8.2.06siy)
- Siyanova-Chanturia, A., & Pellicer-Sánchez, A. (2019). Formulaic Language: Setting the scene. In A. Siyanova-Chanturia & A. Pellicer-Sánchez (Eds.), *Understanding formulaic language: A second language acquisition perspective* (pp. 1–16). New York: Routledge.

- Siyanova-Chanturia, A., & van Lancker, D. S.** (2019). What online processing tell us about formulaic language. In A. Siyanova-Chanturia & A. Pellicer-Sánchez (Eds.), *Understanding formulaic language: A second language acquisition perspective* (pp. 38–61). New York: Routledge.
- Sonbul, S., & Schmitt, N.** (2013). Explicit and Implicit lexical knowledge: Acquisition of collocations under different input conditions. *Language Learning*, **63**, 121–159. Doi: [10.1111/j.1467-9922.2012.00730.x](https://doi.org/10.1111/j.1467-9922.2012.00730.x)
- Szudarski, P.** (2012). Effects of meaning- and form-focused instruction on the acquisition of verb–noun collocations in L2 English. *Journal of Second Language Teaching and Research*, **1**, 3–37.
- Szudarski, P., & Carter, R.** (2016). The role of input flood and input enhancement in EFL learners' acquisition of collocations. *International Journal of Applied Linguistics*, **26**(2), 245–265. Doi: [10.1111/ijal.12092](https://doi.org/10.1111/ijal.12092)
- 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. Doi: [10.1111/lang.12343](https://doi.org/10.1111/lang.12343)
- Vilkaitė, L.** (2017). Incidental acquisition of collocations: Effects of adjacency and prior vocabulary knowledge. *ITL – International Journal of Applied Linguistics*, **168**(2). Doi: [10.1075/itl.17005.vil](https://doi.org/10.1075/itl.17005.vil)
- Waring, R., & Takaki, M.** (2003). At what rate do learners learn and retain new vocabulary from reading a graded reader? *Reading in a Foreign Language*, **15**, 130–163.
- Webb, S.** (2007). The effects of repetition on vocabulary knowledge. *Applied Linguistics*, **28**, 46–65. Doi: [10.1093/applin/aml048](https://doi.org/10.1093/applin/aml048)
- Webb, S., & Chang, A.** (2015). Second language vocabulary learning through extensive reading with audio support: How do frequency and distribution of occurrence affect learning? *Language Teaching Research*, **19**, 667–686. Doi: [10.1177/1362168814559800](https://doi.org/10.1177/1362168814559800)
- Webb, S., & Chang, A.** (2020). How does mode of input affect the incidental learning of collocations? *Studies in Second Language Acquisition*, Available in FirstView. Doi: [10.1017/S0272263120000297](https://doi.org/10.1017/S0272263120000297)
- Webb, S., & Kagimoto, E.** (2009). The effects of vocabulary learning on collocation and meaning. *TESOL Quarterly*, **4**, 55–76. Doi: [10.1002/j.1545-7249.2009.tb00227.x](https://doi.org/10.1002/j.1545-7249.2009.tb00227.x)
- Webb, S., & Kagimoto, E.** (2011). Learning collocations: Do the number of collocates, position of the node word, and synonymy affect learning? *Applied Linguistics*, **32**, 259–276. Doi: [10.1093/applin/amq051](https://doi.org/10.1093/applin/amq051)
- Webb, S., Newton, J., & Chang, A.** (2013). Incidental learning of collocation. *Language Learning*, **63**, 91–120. Doi: [10.1111/j.1467-9922.2012.00729.x](https://doi.org/10.1111/j.1467-9922.2012.00729.x)
- Wray, A.** (2002). *Formulaic language and the lexicon*. Cambridge: Cambridge University Press.
- Zahar, R., Cobb, T., & Spada, N.** (2001). Acquiring vocabulary through reading: Effect of frequency and contextual richness. *Canadian Modern Language Review*, **57**, 541–572. Doi: [10.3138/cmlr.57.4.541](https://doi.org/10.3138/cmlr.57.4.541)

Cite this article: Pellicer-Sánchez, A., Siyanova-Chanturia, A., and Parente, F. (2022). The effect of frequency of exposure on the processing and learning of collocations: A comparison of first and second language readers' eye movements. *Applied Psycholinguistics* **43**, 727–756. <https://doi.org/10.1017/S014271642200011X>