Children's L2 vocabulary acquisition through extensive reading

Brittney Kalogeros - University of Barcelona
brittneykalogeros@gmail.com
Raquel Serrano - University of Barcelona
raquelserrano@ub.edu
Elsa Tragant - University of Barcelona
tragant@ub.edu

Rebut / Received: 15-01-22
Acceptat / Accepted: 21-06-22

Resum. Adquisició de vocabulari infantil L2 mitjançant la lectura extensiva. Estudis anteriors han demostrat que la lectura extensiva (LE) fomenta l'aprenentatge del vocabulari tant de L1 com de L2. L'estudi actual investiga més el potencial de la LE en el cas dels alumnes de primària. Vint-i-tres estudiants d'EFL d'una aula intacta de 4t (9-10 anys) van participar en deu sessions de LE amb suport d'àudio. Els estudiants van realitzar una prova de vocabulari que inclou 25 paraules abans, immediatament després i tres mesos després del programa de LE. Els guanys de vocabulari de L2 es van analitzar segons la freqüència de les paraules (nombre d'exposicions a les paraules meta) i la distribució (el nombre de llibres en què apareixien les paraules meta). Els resultats mostren que la LE va provocar guanys significatius de vocabulari. A més, es van trobar efectes significatius de la freqüència i la distribució en els guanys de vocabulari immediats i de la freqüència en els guanys a llarg termini.

Paraules clau: *aprenentatge incidental de vocabulari, lectura extensiva, freqüència, distribució.*

Abstract. Children's L2 vocabulary acquisition through extensive reading. Previous studies have shown that extensive reading (ER) encourages both L1 and L2 vocabulary learning. The current study further investigates the potential of ER in the case of primary school learners. Twenty-three Spanish EFL learners from an intact grade 4 classroom (9-10 years old) participated in ten ER sessions with audio support. The learners performed a vocabulary test including 25 target words before, immediately after and three months after the ER programme. L2 vocabulary

gains were analysed according to word frequency (number of exposures to the target words) and distribution (the number of books in which the target words appeared). Results show that ER led to significant vocabulary gains. Additionally, significant effects of frequency and distribution were found on immediate vocabulary gains and of frequency on long-term gains.

Keywords: *incidental vocabulary learning, extensive reading, word frequency, distribution.*

Introduction

Extensive reading (ER), a practice involving reading many, often easier, texts over a longer period of time, has been found beneficial to develop L2 learners' vocabulary. By reading extensively, L2 learners are exposed to large amounts of input, which helps them fill in holes in their linguistic repertoire and make form-meaning connections that can eventually lead to acquisition (Webb & Chang, 2015). Research into incidental vocabulary learning through ER has been a topic of great interest for researchers (Cho & Krashen, 1994). However, it is still unclear to what extent EFL learners, especially young learners, can acquire vocabulary from reading children's graded readers. These adapted materials are especially produced with English learners in mind and their multimodal nature (pictures, text and very often audio) makes them quite suitable for young L2 learners (Kaminski, 2019). The current study contributes to this body of literature by examining how a group of learners acquired new vocabulary through ER of the graded readers of their choice over several sessions implemented during their EFL class. This study also investigates the role of word frequency (how many times a word is encountered) and distribution (over how many books) on learners' vocabulary acquisition.

1. Literature Review

1.1. Incidental vocabulary acquisition through ER

ER is a popular pedagogical activity for both L1 and L2 learners which is related to pleasure reading. In an L2 classroom context, ER often entails students regularly reading full books, or graded readers, for about 30 minutes to an hour per session. According to Brown, Waring and Donkaewbua (2008), readers should ideally be familiar with the majority of vocabulary included in the books, and will be granted the freedom of selecting their own stories to read, while reading at their own pace and at a level they can understand (for an updated and thorough review of ER in the L2 class, see Nation & Waring, 2020). ER is a good complement to explicit teaching because it allows for opportunities for vocabulary development that explicit teaching cannot provide, given the large amount of vocabulary that is needed to deal with most spoken (6000-7000

word families) and written texts (8000-9000 word families) (Nation, 2006). Through ER, learners can fill in gaps in their vocabulary knowledge, which differ between individuals, while enjoying the reading process. This practice encourages learners to read faster and more, resulting in an increase in comprehension, circling back to an increase in the difficulty of text; a process coined the "virtuous circle of the good reader" (Nuttall, 1996, p. 127). ER programmes have been claimed to lead to a robust development of vocabulary and reading comprehension skills in an ecological way for learners of all age groups (e.g., Elley & Mangubhai, 1983), even though there is a scarcity of research with children in ER interventions (Nakanishi, 2015).

Research shows that incidental vocabulary learning is crucial in both L1 and L2 learning processes (Nation, 2001). The present study operationalizes incidental vocabulary learning as "a by-product" of an activity that does not include an explicit vocabulary focus, in this case, reading (Hulstijn, 2003).

ER is one of the central activities to Nation's (2001) strand of meaning-focused input. Many studies have reported general benefits of such input, especially with respect to incidental vocabulary learning (Pigada & Schmitt, 2006). For instance, Suk (2017) examined the impact of including a 15-week quasi-experimental ER intervention as a curriculum component within an existing curriculum for a group of Korean university EFL learners. The results of the study confirmed the benefit of integrating ER into the existing curriculum, as the students who participated in the programme significantly outperformed their classmates in reading comprehension, reading rate, and especially vocabulary knowledge

In the case of children, participation in ER programmes also seems to be beneficial, although research with this population is scarce. Lightbown and colleagues (Lightbown, 1992; Lightbown, Halter, White & Horst, 2002) found that a group of Canadian students who learned English by reading-while-listening to materials of their choice obtained similar L2-learning results as children who had received teacher-led instruction through an audio-lingual approach. Similar findings were obtained in a smaller-scale study (Tragant, Muñoz & Spada, 2016) conducted in Spain where school-aged learners who were exposed to extensive reading 60% of their EFL instruction time showed comparable results to a group who received teacher-led instruction only. These studies focused on different L2 areas and learners' motivations, but did not provide a thorough analysis of learners' vocabulary improvement in relation to the books that they read, which is the main aim of the current paper.

1.2. Word Frequency and Distribution in L2 Vocabulary Acquisition through Reading and ER

A number of factors have been identified as predictors of incidental vocabulary learning through reading and frequency seems to be one of the strongest predictors (Vidal, 2011). In the case of ER in particular, Nation and Waring (2020) claim that "vocabulary learning from extensive reading occurs as a result of repeated meetings (our italics) with words, word

families, and lexical phrases in context, and as a result of the quality of these meetings." (p. 72). Indeed, word frequency explains 11 % of the variance in word learning through input, according to Uchihara, Webb and Yanagisawa (2019). Despite the number of studies on the effects of word frequency on incidental vocabulary learning from reading, it is difficult to determine how many times a word should be encountered before it is learned, as there are many other factors besides frequency that can affect word learning (for a review see Peters, 2019). Several authors, however, have provided different suggestions as to the recommended number of encounters with novel words. For instance, according to Chen and Truscott (2010) seven encounters are needed for a discernible rise in learning rate whereas according to Brown et al. (2008), 15-20 encounters are necessary. Another factor that could affect vocabulary learning through reading is how repeated exposures to new words are distributed across texts or books. The studies that have examined this provide conflicting evidence, with some studies suggesting a benefit for shorter inter-repetition lags for incidental vocabulary learning (Elgort & Warren, 2004), while others find no differences (Webb & Chang, 2015) or differential effects for word form and meaning (Elgort et al., 2018). More research is necessary to throw more light on the topics of frequency and distribution both for theoretical as well as practical purposes.

In their ER intervention including graded readers, Webb and Chang (2015) investigated incidental vocabulary learning through extensive reading-while-listening, as well as the effect of frequency and distribution of exposures to target words on vocabulary gains. The ER programme was implemented with a group of 61 EFL Taiwanese learners (aged 15-16), who read and listened to the same ten graded readers. The students performed a bilingual matching vocabulary test before and after the ER programme, which included 100 words that were randomly selected from the graded readers. The authors reported relative vocabulary gains of 44.06 % immediately after the programme and 36.66 % three months later. However, frequency and distribution of exposures to the target words were not found to have a significant relationship with vocabulary gains, contrary to the authors' prediction. They attributed this outcome to a number of elements in their design, such as including too few books (ten) with a small range of topics and too many target words (100). The authors concluded that frequency was one of many factors that affected vocabulary acquisition in their study and encouraged more research on the topic. Additionally, it should be pointed out that the reading programme in Webb and Chang (2015) lacks one of the main components of ER, which involves learners' choice of books and reading at their own pace.

2. The Present Study

The current study aims to provide additional evidence for the effect of ER on incidental vocabulary learning by examining an ER programme for primary school children, who are an under-researched population. Moreover, another goal of this study is to further examine the effects of frequency and distribution on incidental vocabulary gains,

extending the findings from Webb and Chang (2015). Our study differs from Webb and Chang (2015) in some important aspects. First, students were younger (9-10 years old), and, as suggested above, very little research exists on how these young L2 learners can acquire vocabulary from reading. Additionally, students had a wider range of books to choose from (n = 34 vs. 10) spanning many topics. Moreover, the ER programme in the present study fostered students' autonomy in choosing their own books and reading at their own pace with individual MP3 players, allowing for a more ecologically valid investigation of the impact of ER. Another important difference between the study by Webb and Chang and the present one is that the classroom teachers participated in the selection of the target words in the current study, adding to the validity of the choice. The specific research questions that guide this study are the following:

1. To what extent can ER with audio support promote incidental vocabulary learning in the case of young EFL learners?
2. Is frequency of exposure to new words a determining factor in vocabulary gains through ER?
3. Does the distribution of repeated encounters to new words in different books affect incidental vocabulary learning?

3. Methodology

3.1. Participants

The participants in this study were 23 (13 male and 10 female) Catalan/Spanish speaking 9-10-year-old students from an intact classroom (grade 4). At the time of data collection, the learners had accumulated 1,900 hours of English instruction in total and their English proficiency was estimated as A1 (CEFR 2001). The ER programme was implemented by their teacher, who was an experienced EFL teacher with previous experience with ER.

3.2. Materials and design

3.2.1. Graded Readers

The graded readers included 34 titles which were selected by the teacher (from a pool of 94 titles provided by the researchers) to be given to students based on what best matched the students' reading abilities. The graded readers belonged to various series from six different publishers and were addressed to primary school learners. The books were accompanied by commercially produced audio narration of the story (length ranging from 5 to 23 minutes) as part of the reader's set. Students read an average of 16 books each, with a range of 12-23 books read per learner.

Although the students could choose the books that they liked, there were four graded readers that all students read, to make sure there was a list of target words that all the students would have seen at least once. These four graded readers were from the "Uncle Jack" series, published by Young ELI readers. They were at the A1.1 level (CEFR 2001) and most of the words (84 %) were at the K1 word level. These graded readers were adventure stories based on the same characters about different environmental issues such as deforestation, illegal hunting in Africa etc.

3.2.2. Target Words and Vocabulary Tests

From the four Uncle Jack graded reader titles, 25 words were selected as target items for vocabulary tests, although many of these words also appeared in other books the children read. The target words included 2 verbs, 5 adverbs, 5 adjectives and 13 nouns. The 25 target words were made up of 12 words from the first 1,000 most frequent words in English (K1 word level), 8 words from the first 2,000 (K2 level), one word each from the K3, K4, K5 and K6 word levels. There was one word from which we cannot report lexical frequency because one of them was a two-word expression ('at once'). The words were selected from a list provided to two of the school's EFL teachers as possible target words as they were thought to be predominantly unknown by students. Considering students read their own selection of the provided graded readers, each of them had different numbers of encounters with the target words. Table 1 presents the list of the target words included in the books the students chose in alphabetical order, together with the range of encounters that students experienced (minimum and maximum frequency) and the range in number of different books in which they appeared to the learners (minimum and maximum distribution).

Table 1. List of target words and range of encounters in frequency and distribution

Target Word	Minimum Number of Frequency Encounters	Maximum Number of Frequency Encounters	Minimum Distribution Encounters	Maximum Distribution Encounters
ant(s)	2	6	1	2
at once	1	2	1	2
blanket(s)	1	1	1	1
cage(s)	5	25	1	3
crazy	1	1	1	1
curly	1	1	1	1
ground	8	17	3	9
hill(s)	2	10	1	4

hole(s)	5	15	1	5
hug	1	1	1	1
jug	1	1	1	1
land	1	2	1	2
loudly	1	4	1	4
maybe	7	22	2	6
muddy	1	1	1	1
nearly	1	5	1	4
net(s)	1	1	1	1
nut(s)	4	4	1	1
quietly	1	3	1	3
ride(s)	1	2	1	2
saw(s)	1	1	1	1
smelly	1	2	1	2
strip(s)	3	3	1	1
twin(s)	3	3	1	1
weak	1	1	1	1

Target word knowledge was assessed with a bilingual L2-L1 matching test targeting meaning recognition, as in Webb and Chang (2015). The format of the test included the English target word on the left side of the page to be matched to the Catalan/Spanish definition on the right side of the page. The items were presented in five groups of five words. There was one distractor included in each group among the possible Catalan/Spanish translations (see Figure 1 for an example of a set of items).

1. weak	boig/ loco ____
2. smelly	malhumorat (que sovint es queixa) / gruñón ___
3. crazy	pudent (que fa pudor) / que apesta ___
4- curly	dèbil / débil
5. muddy	enfangat / fangoso (lleno de barro) ___
	arrissat (cabell) / rizado ___

Figure 1. Examples of items from L2-L1 matching test

4. Procedure

4.1. ER treatment

The intervention was carried out over ten sessions, two per week (45 minutes and 1 hour respectively). At the beginning of the first session, the students were instructed to look through the books, at the pictures and at the glossary before beginning to read to help them get an idea of the subject of the book. They were also instructed to start with books that they found easy to read. Each student was given an MP3 player to support the reading with audio and read at their own pace.

At the beginning of each session, students chose the books they would read. The cover of the book displayed the level of difficulty and the length of the audio, so they could make informed decisions about their selection. After reading each book, the students completed a worksheet about the content of the book (i. e. "My favourite character was…") as well as a question about which words from the book they would like to learn. The teacher was asked to support students as needed during independent reading, while completing the worksheets, when selecting the books, and in operating the MP3s.

Testing

In order to measure students' prior knowledge of the target words, the students performed the vocabulary pre-test two days prior to the intervention. Three days after the final reading session, the students took the post-test to measure if/which target words were acquired during the treatment (immediate gains). Three months after the first post-test, the same test was administered as a delayed post-test to examine long-term gains. All the tests were unannounced and they took an average of 12 minutes to complete.

5. Analyses

5.1. Vocabulary knowledge

Each vocabulary item was coded independently as '0' when the students' response on the multiple-choice test was incorrect and '1' when it was correct. These scores by item were used to compare students' vocabulary scores across the three testing times.

In order to examine how vocabulary gains were related to frequency and distribution, we calculated vocabulary gains scores by items in the following way. We coded learning gains as '0' when the pre-test score was '0' or '1' and the post-test score was '0'. We coded learning gains as '1' when the pre-test score was '0' and the post-test score was '1'. Items that had a score of '1' both at pre-test and post-test were excluded from the analysis, as they were known words. The same coding procedure was followed to examine long-term gains using the scores of the pre-test and delayed post-test. Additionally, relative gains

were also calculated to examine percentage gain considering pre-test scores by using the following formula: [(post-test/delayed post-test score - pre-test score) / (1 - pre-test score)] x 100.

5.2. Word frequency and distribution

Frequency was examined by counting the number of times the target words appeared in the books the children read. As not all the children read the same books, frequencies were different for the target words. The 25 target words were encountered in a frequency range from 1-25 times (Table 1). In total, there were 572 encounters of the target words throughout all the books, considering all the students together. Four different frequency groups were created for the analyses, from least to most frequent. Group 1 included the words that students were exposed to once (273 cases), Group 2 consisted of the words that appeared two to four times (187 cases), Group 3 included words appearing five to ten times (60 cases), and finally Group 4 included words appearing more than ten times (52 cases).

Finally, we analysed distribution considering the number of books in which the target words appeared. As seen in Table 1, in most cases, words were repeated within the same book, but the distribution range was from one to nine books. Three groups were made: Group 1 included the words that appeared in only one book (364 cases), Group 2 from two to three books (158 cases), and Group 3 included the words appearing in more than three books (50 cases). The frequency as well as the distribution groups were made considering the data that we had, trying to make groups that were as balanced as possible in terms of number of cases and number of words included within each group.

5.3. Statistical analyses

The statistical analyses were computed using SPSS v25. In order to answer the first research question and compare the students' knowledge of the target words before and after the ER intervention, we performed two related-sample *t*-tests, one with the scores of the pre- and post-test to< examine immediate gains, and another one with the pre- and delayed post-test scores to analyze long-term gains. To answer the second and third research questions, a series of Generalized Linear Mixed Models (GLMMs) were conducted and Bonferroni pairwise comparisons calculated. All models included subject and item as random factors. The fixed effects of word frequency and word distribution were analyzed separately because, as expected, there was a significant correlation between the number of word repetitions (frequency) and the number of books where the words appeared (distribution): the higher the frequency, the higher the distribution as well.

6. Results

6.1. Vocabulary gains

Table 2 presents the descriptive means and standard deviation per item for each testing time, out of a possible maximum score of 1. It can be observed that both post- and delayed post-test scores were higher than pre-test scores.

Table 2. Descriptive statistics of vocabulary scores (maximum 1)

	Mean	**SD**
Pre-test	.34	.47
Post-test	.47	.50
Delayed post-test	.41	.49

The results of the *t*-tests suggest that there were statistically significant differences between pre-test and immediate post-test scores: $t(571) = -5.35$, $p < .001$, Cohen's $d = 0.45$ and also between pre-test and delayed post-test scores: $t(571) = -3.09$, $p = .002$, Cohen's $d = 0.26$. This indicates that the ER treatment had promoted significant immediate and long-term incidental vocabulary gains. In terms of relative gains considering pre-test scores (i.e. how much students could learn considering what they already knew), there was a 19.70% immediate and 10.60 % long-term gain.

6.2. Effects of frequency on vocabulary gains

Figure 2 presents the immediate vocabulary gains for each frequency group out of a maximum score of 1. It can be seen that, as frequency increased, vocabulary gains also increased.

The results of the statistical analyses suggest that the effect of frequency was statistically significant: $F(3, 430) = 2.69$, $p = .046$. The Bonferroni pairwise comparisons indicate that after ten exposures differences started becoming significant compared to the lower frequencies (small effect size). See results of all pairwise comparisons in the Appendix. The scores on the delayed post-test follow a similar pattern, suggesting that more frequency led to higher long-term gains (Figure 3).

The effect of frequency was also statistically significant: $F(3, 438) = 2.69$, $p = .046$ on long-term gains. The pairwise contrasts show that the only significant contrast (with a small effect size) was between one exposure versus ten or more (see Appendix 1).

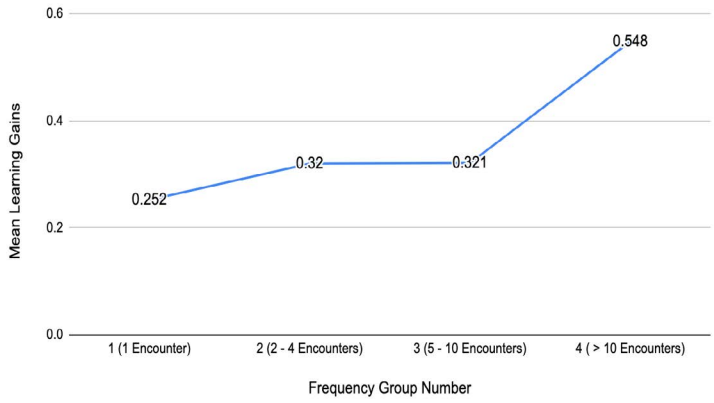

Figure 2. Frequency effects on immediate gains (maximum = 1)

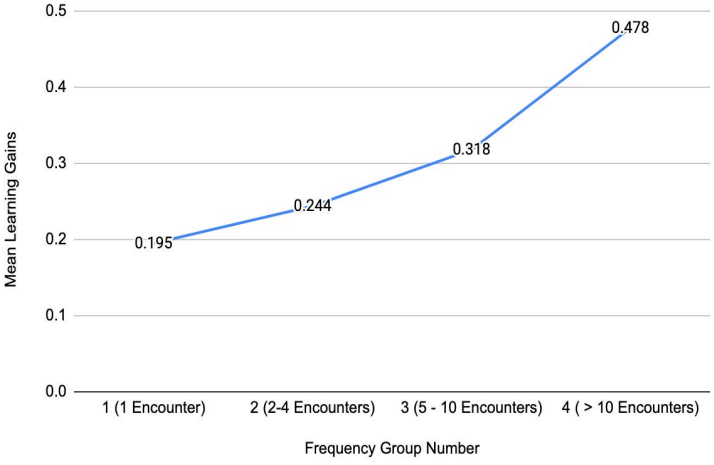

Figure 3. Frequency effects on long-term gains

6.3. Effects of distribution on vocabulary gains

In the case of distribution, there was not a linear increase in immediate gains as the number of books where the target words appeared also increased. Results show that the highest gains were obtained when the words appeared in two to three books (Figure 4).

The effect of distribution was approaching significance: $F(2, 431) = 2.90$, $p = .056$. The results of the pairwise comparisons show that there was a statistically significant

difference (small effect size) between the words that appeared in one book versus two-three books, the latter group leading to higher vocabulary gains.

In the case of long-term gains, Figure 5 shows a slightly different pattern, in the sense that a larger distribution led to higher long-term vocabulary gains. However, the results of the analyses show that the effect of distribution in this case was not statistically significant: $F(2, 439) = 1.50$, $p = .225$.

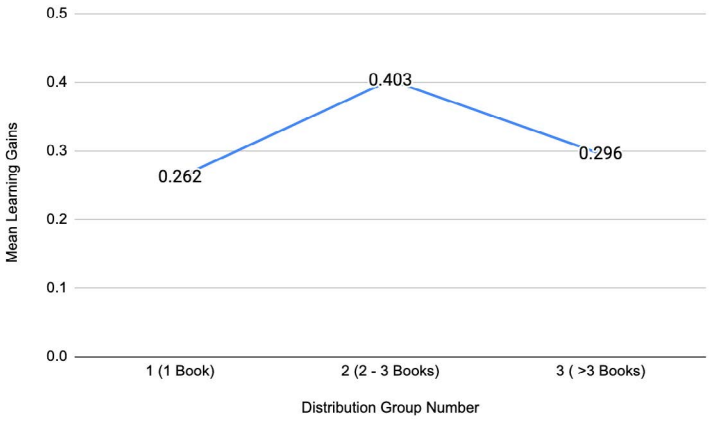

Figure 4. Distribution effects on immediate gains

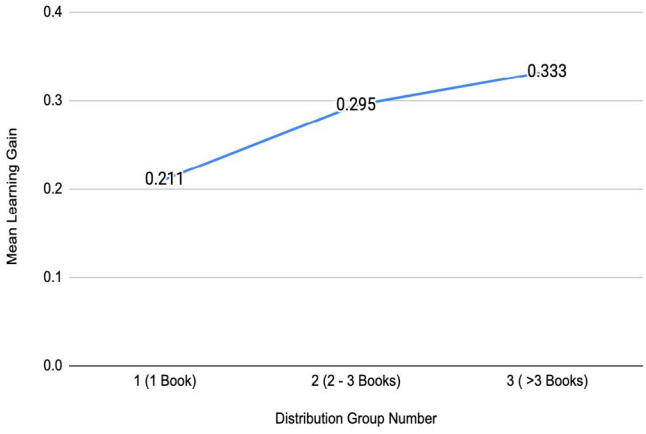

Figure 5. Distribution effects on long-term gains

7. Discussion

The current study examined incidental vocabulary gains through ER and whether frequency and distribution of encounters in the graded readers affected L2 vocabulary learning. With respect to the first research question, our results suggest that the ER programme promoted immediate and long-term vocabulary learning. When examining the overall gains, they were lower than in Webb and Chang (2015), but in line with other studies on incidental vocabulary learning from reading with children (Tragant, Llanes & Pinyana, 2019). The higher relative gains reported by Webb and Chang could be related to the older age of the learners as well as to the fact that words were repeated more often. Whereas in Webb and Chang (2015) 37 % of the words appeared more than 10 times, in the current study, less than 10 % of the 572 exposures to the target words included words appearing more than 10 times. Additionally, the Taiwanese students got involved in discussions about the readings as well as reading aloud segments of the book, which could also explain the differences in results between the study by Webb and Chang and the present study, which did not involve any group or whole-class post-reading activities.

Regarding the second research question, our results suggest that frequency of exposure to novel words had a significant effect on both immediate and long-term incidental vocabulary learning. When words were encountered more than ten times they were better learned and remembered, which is in line with other previous studies (e. g., Pellicer-Sánchez & Schmitt, 2010; Webb, 2007). As previously stated, there is no agreement in the literature as to how many encounters are necessary for words to be learnt incidentally from reading. Our results with children are in line with those in Webb (2007) with university students, who found that after 10 encounters with target words sizable gains occurred. On the other hand, Webb and Chang (2015) found no significant effects of frequency and assert that this could be due to the numerous additional factors in their treatment that could also affect learning (discussion of the stories, use of dictionaries, etc.)

In terms of distribution, Webb and Chang (2015) did not find a significant effect for this variable on vocabulary learning. The present study confirms this finding, but only for long-term gains. In the case of immediate gains, we found that repeated exposures over two to three books led to more learning than when words were repeated within the same book. However, distribution over four or more books did not lead to additional vocabulary gains. The effect of distribution was only approaching significance. However, this finding is in line with some claims in the literature that spacing repetitions of novel words is beneficial unless spacing is so wide that previous encounters with the target words are no longer remembered (Toppino & Bloom, 2002). This might have been the case when words were distributed over many books (4-9). Our results for distribution should be interpreted cautiously as the overwhelming majority of the words were repeated within the same book and there were very few cases (around 8 %) in which words appeared in more than three books. Additionally, we could not control for

the temporal spacing between readings of different books where the target words were repeated.

8. Conclusions and Pedagogical Implications

Incidental vocabulary learning is closely associated with ER programmes where students have the option of choosing what books they read (Day & Bamford, 1998). Yet, it is often the case that research on incidental vocabulary learning through reading does not often give learners this option (Webb & Chang, 2015). This can be explained by the fact that having all learners read the same text greatly simplifies the task of measuring word learning. In the present classroom-based study, our students were given this choice and yet we found a way to still be able to measure incidental vocabulary gains. Future studies should include a higher number of words and also other vocabulary tests that might measure dimensions of vocabulary knowledge other than meaning recognition, including form and meaning recall and form recognition. Additionally, it should be acknowledged that, since the participants did the same vocabulary test multiple times, this test-repetition effect could have partly affected the results. Future studies could control for this, by having only half of the participants take the immediate post-test and the other half the delayed post-test. It must also be mentioned that, although our results show some statistically significant differences, the effect sizes were not large, which also points to the importance of doing more studies on this topic.

Our results suggest that ER is overall recommendable as a classroom practice with children to enhance vocabulary learning in the L2, as shown by the statistically significant vocabulary gains experienced by the learners. Although the learning gains were moderate, it must be kept in mind that learners in the present study only participated in ten ER sessions and that more than half of the target words were encountered only one to three times. The results also suggest that increased exposure to the same words through reading leads to higher gains, which are especially significant after ten exposures. Considering this, it might be recommendable to choose graded readers from the same series or similar topics, as they probably encourage repetition of a set of common words. Alternatively, repeated readings of the same books might also be recommended as a way to promote increased exposure to the same target words. In the case of repetitions of words across books, it is less clear how words should be ideally distributed, but our results seem to suggest that it is better that repeated exposures are distributed over more than one book.

In conclusion, the results of our study present additional evidence for the positive effects of ER with children and encourage the implementation of this practice in EFL classrooms in conditions that promote repeated exposures to L2 words in order to foster their incidental learning.

References

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

Chen, C., y Truscott, J. (2010). The effects of repetition and L1 lexicalization on incidental vocabulary acquisition. *Applied Linguistics*, *3*, 693-713.

Cho, K., & Krashen, S. (1994). Acquisition of vocabulary from the sweet valley kids series: Adult ESL acquisition. *Journal of Reading, 37*(5), 662-667.

Day, R., & Bamford, J. (1998). Extensive reading in the second language classroom. *Cambridge University Press*.

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.

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.

Elley, W. B., & Mangubhai, F. (1983). The impact of reading on second language learning. *Reading Research Quarterly, 19*(1), 53-67.

Hulstijn, J. H. (2003). Intentional learning". In C. Doughty & M. Long (Eds.), The Handbook of Second Language Acquisition. (p. 349-381). Blackwell Publishing Ltd.

Kaminski A. (2019). Young learners' engagement with multimodal texts. *ELT Journal*, *73*(2), 175-185.

Lightbown, P. M. (1992). "Can they do it themselves? A comprehension-based ESL course for young children". In R. Courchene, J. St. John, C. Therrien, & J. Glidden (Eds.), *Comprehension-based second language teaching* (p. 353-370). University of Ottawa Press.

Lightbown, P. M., Halter, R. H., White, J., & Horst, M. (2002). Comprehension-based learning: The limits of "do it yourself". *Canadian Modern Language Review/La Revue canadienne des langues vivantes, 58*(3), 427-464.

Nakanishi, T. (2015). A meta-analysis of extensive reading research. *TESOL Quarterly*, *49*(1), 6-37. http://dx.doi.org/10.1002/tesq.

Nation, I. S. P. (2001). Learning Vocabulary in Another Language. *Cambridge University Press.*

Nation, I. S. P. (2006). How large a vocabulary is needed for reading and listening. *The Canadian Modern language Review*, *63*(1), 59-82.

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

Nuttall, H. C. (1996). Teaching reading skills in a foreign language (2nd Edition). *ELT Journal, 52*(2), 169-173.

Pellicer-Sánchez, A. &Schmitt, N. (2010). Incidental vocabulary acquisition from an authentic novel: Do *Things Fall Apart*? *Reading in a Foreign Language, 22*(1), 31-55.

Peters, E. (2019). Factors affecting the learning of single-word items. In S. Webb (Ed.), *The Routledge handbook of vocabulary studies*,125-142. Routledge.

Pigada, M. & Schmitt, N. (2006). Vocabulary acquisition from extensive reading: A case study. *Reading in a Foreign Language*, *18*(1), 1-28.

Suk, N. (2017). The effects of extensive reading on reading comprehension, reading rate, and vocabulary acquisition. *Reading Research Quarterly, 52*(1), 73-89.

Toppino, T. C. & Bloom, L. C. (2002). The spacing effect, free recall, and two-process theory: A closer look. *Journal of Experimental Psychology: Learning, Memory and Cognition, 28*(3), 437-444.

Tragant, E., Muñoz, C., & Spada, N. (2016). Maximizing young learners' input: An intervention program. *Canadian Modern Language Review, 72*(2), 234-257.

Tragant, E., Llanes, À. & Pinyana, À. (2019). Linguistic and non-linguistic outcomes of a reading-while-listening program for young learners of English. *Reading and Writing, 32*(3), 819-838.

Uchihara, T., Webb, S. & Yanagisawa, A. (2019). The effects of repetition on incidental vocabulary learning: A meta-analysis of correlational studies. *Language Learning*, *69*(3), 559-599. https://doi.org/10.1111/lang.12343.

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

Webb, S. (2007). The effects of repetition on vocabulary knowledge. *Applied Linguistics, 28*(1), 46-65.

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*(6), 667-686.

Appendix

Pairwise contrasts (statistically significant results in bold)

	Pair 1	**Pair 2**	***t***	***df***	***p***	**Cohen's *d***
Immediate Gains in Frequency Groups	1	2-4	-1.220	430	.223	
		5-10	-.759	430	.448	
		+10	-2.705	430	**.007**	0.26
	2 - 4	5-10	-.009	430	.993	
		+10	-2.018	430	**.044**	0.19
	5 - 10	+10	-1.771	430	.077	
Long-term Gains in Frequency Groups	1	2-4	-.929	438	.354	
		5-10	-1.333	438	.183	
		+10	-2.475	438	**.014**	0.23
	2 - 4	5-10	-.786	438	.433	
		+10	-2.002	438	**.046**	0.19
	5 - 10	+10	-1.201	438	.230	
Immediate Gains in Distribution Groups	1	2-3	-2.327	431	**.020**	0.22
		+3	-.367	431	.714	
	2 - 3	+3	1.092	431	.275	