

# Extensive reading of picture books and incidental vocabulary learning: a comparison between false beginners and more advanced learners<sup>1</sup>

Naomi KIHARA

## 要 約

本稿は英語多読におけるL2学習者の語彙習得に関して、特に、擬似初心者と習熟度が高い学習者の比較を中心に考察を行ったものである。英語多読図書としてOxford Reading Trees (ORT) シリーズ(ステージ1～5の計154冊)を使用し、大学生9名に多読指導を半年間行い、①筆者が指定した45の英単語の和訳テスト、②個々の学生が多読中に作成した個人別英単語の和訳テスト、を実施した。①のテストについては、ORT154冊のテキストを全てデジタル化してコーパスを作成し、頻度別に単語45個を任意に抽出したものを多読開始直前と直後に実施した。②については多読中に学生が自発的に記録した単語学習リストを基に個別の単語テストを作成したものであり、多読の全過程終了後に実施した。考察の結果、多読図書において頻度の高い単語の方が低い単語よりも習得される確率が高いことが示唆されたが、その傾向は習熟度が高い学生に顕著であり、擬似初心者の単語習得は頻度にかかわらず限定的であることがわかった。また、ほとんどの学生が多読を通して自発的な未知語の学習を行っていたが、学習した単語の数、多読終了後の定着数ともに習熟度が高い学生の方が高く、擬似初心者は多読終了時点において単語の定着数が低いことが明らかになった。

## 1 Introduction

Recent research into extensive reading (ER) in EFL teaching and learning has drawn attention to issues relating to the acquisition of vocabulary through extensive reading, with a particular focus on incidental learning (for example, Brown, Waring and Donkaewbua, 2008; and Kweon and Kim, 2008). The author's previous study (Kihara, 2009a) also suggested the possibility that carefully selected extensive reading books (picture books with very easy English) could be beneficial in vocabulary acquisition for false beginners who have limited vocabulary and grammatical knowledge. In this study, 17 students whose English proficiency was at the false beginner level read a total of from 67,000 to 200,000 words in one year. When interviewed (Kihara, 2008), all of the 17 students stated that they were able to guess unknown words during reading by looking at pictures in the books. The guessing processes of some of the students were observed by the author during think-aloud interviews. Despite finding evidence for some incidental learning of unknown words in some students, however, the author was

---

<sup>1</sup> A preliminary report on this study was presented at the 2009 PKETA International Conference in Busan, Korea (Kihara, 2009b).

unable to conclude to what extent ER is beneficial in vocabulary development for false beginners. In order to explore the relationships between extensive reading of picture books and vocabulary development, this paper investigates the following research questions:

- 1) After reading a certain amount of ER books, does a student show any improvement in his/her understanding of the words used in the ER texts compared with understanding before reading? If so, is the frequency of a word in the ER texts related to the improvement in word understanding?
- 2) When encountering an unknown word during ER, does a student make a correct guess as to its meaning? Or does a student attempt to learn the meaning of the word by referring to a dictionary? If so, are the meanings of the words remembered after ER?
- 3) How do these results differ between false beginners and non-false beginners?

## **2 Methodology**

### **2.1 Subjects**

Nine university students were engaged in this study. Four of them had an English proficiency at the false beginner level, two were low-intermediate, and three were at a more advanced level. All of the three more advanced students were Chinese speakers and the rest were Japanese speakers. The students met once a week for 14 weeks for a 90 minute-'Extensive Reading' class and spent 45-60 minutes for ER each class.

### **2.2 Books for ER**

All of the participant students were asked to read books from the Oxford Reading Tree (ORT) series (from stage 1 to stage 5, a total of 154 books) which is an authentic material that is designed for English native speaker children. While the text of each book is extremely short, with the most basic vocabulary and sentence structures, each text element is accompanied by a clear picture that replicates the linguistic content of that element in visual form. For this reason, these texts would appear to present a significant opportunity for EFL students, even those who are at the false beginner level, to guess the meanings of previously unknown words and to expand their vocabulary knowledge. Originally, the students were also asked to read the Longman Literacy Land (LLL) series (from step 1 to step 6, a total of 81 books) after finishing the ORT books, but this instruction was aborted for two reasons. One was that those who were at a lower level of proficiency found reading all of the ORT books (from stage 1 to stage 5) challenging and by the time they finished the whole series, it seemed that the students possessed neither the motivation nor the time to complete the LLL series (from step 1 to step 6) in the allocated time. At the same time, those who were at a more advanced level found reading all of the ORT books very easy and enjoyable, but

were eager to start reading more challenging books than the LLL series. In order to accommodate these students' needs and secure the motivation of reading, the LLL series was withdrawn from the original plan.

## 2.3 Vocabulary tests

### 2.3.1 Testing target vocabulary before and after reading the ORT books

The texts of the ORT and the LLL books were digitalized and analyzed to produce a corpus and subsequently a word frequency list. A list of 45 target words was then generated and used as the basis for meaning recognition (translation) tests. Five words were randomly taken from 9 different frequency bands, as shown in Table 1. Students were given a meaning recognition test before they started to read the ORT series (pre-test), and the same test was re-administered after they had finished reading the series (post-test). As the original plan was to read both the ORT and the LLL series, the texts of the LLL series were included in this corpus. The set of 45 words in Table 1 was given to the students as pre- and post-tests. However, the frequency bands were readjusted as shown in Table 2, in order to represent the word occurrence in the ORT series (not including LLL) for the purposes of analysis. In this paper, 'word frequency' refers to the word occurrence in the ORT corpus unless specifically described otherwise.

Table 1. Target words (arranged in terms of ORT-LLL frequency)

Frequency	Words				
18	things	hole	mess	rope	girl
16	balloon	head	elephant	dream	hat
14	garden	cream	duck	band	knight
12	hens	hose	pond	fight	scarf
10	sandcastle	troll	bath	lorry	mud
8	crane	thorns	pirate	vet	wallpaper
6	tap	ladder	telescope	volcano	rubbish
4	flour	twins	tail	hill	hutch
2	fist	seals	giraffes	hay	harness

Table 2. Target words (re-arranged in terms of ORT frequency)

Frequency	Words							
14	balloon							
13	mess	elephant	duck					
12	girl							
11	rope	cream	band	hose	scarf			
10	things	hole	fight	troll	lorry			
9	hat	knight	mud					
8	pond	thorns	pirate	wallpaper				
7	bath	crane						
6	tap	volcano						
5	head	sandcastle	ladder					
4								
3	garden							
2	vet	tail	seals	giraffes	harness			
1	dream	hens						
0	telescope	rubbish	flour	twins	hill	hutch	fist	hay

### 2.3.2 Monitoring vocabulary learned during ER and testing the retention rates of words learned immediately after ER

Students were asked to take notes during their ER, when making a guess about an unknown word or when referring to a dictionary. The notes were collected at the end of every ER session. When all of the 14 ER sessions were completed, a tailored word list which presented all the words a student guessed or referred to in a dictionary was produced for each student. The students were given these individually produced word lists as a meaning recognition (translation) test to test how well they remembered the words after finishing the complete ER program.

## 3 Results and discussion

### 3.1 Testing vocabulary before and after reading the ORT books

#### 3.1.1 Score comparison between the pre- test and the post-test

The results of the recognition test on the 45 target words (Table 1) are shown in Table 3. All of the students improved in their number of correct answers. The number of wrong answers, however, also increased at the same time. This means that students' attempts increased (ie. the number of blank answers decreased) even though some of their attempts were futile.

Table 3. Results of pre- test and post-test

Student	Pre-test			Post-test			Improvement from Pre-test to Post test			
	Correct answer	Wrong answer	No answer	Correct answer	Wrong answer	No answer	Correct answer	Average	Wrong answer	No answer
F1	23	3	19	31	4	10	+8	+6.0	+1	-9
F2	18	5	22	24	7	14	+6		+2	-8
F3	18	8	19	21	4	20	+3		-4	+1
F4	15	2	28	22	4	19	+7		+2	-9
LI1	21	5	19	24	5	16	+3	+5.0	0	-3
LI2	21	6	18	28	8	9	+7		+2	-9
A1	25	7	13	31	4	10	+6	+10.3	-3	-3
A2	22	7	16	41	1	3	+19		-6	-13
A3	26	4	15	32	5	8	+6		+1	-7

\*F = false beginner student, LI =low-intermediate student, A = more advanced student

When looking at the average of each proficiency group's improvement, the group of more advanced students (A1, A2, and A3) showed a considerably higher improvement (+10.3) than the low- intermediate (+5.0) and false beginner students (+6.0). These figures must, however, be treated with caution, as the average of the advanced group was boosted by the score of one student (A2), and the rest of the group's (A1 and A3) scores were actually the same as or lower than F1, F2 and F4 in the false beginner group. Although it must be noted that the greatest improvement by far is found in the advanced group and the smallest improvement is seen in the false beginner and the lower-intermediate groups, the improvement of the rest of the students does not seem to show a significant difference (sd of improvement of

correct answers=1.83, excluding A2).

### 3.1.2 Word recognition vs. word frequency

The comparison of the raw scores between the two tests suggests that all of the students improved in recognising words correctly. How is their improvement related to word frequency in the texts they read in ER? Table 4 lists the words in terms of word frequency in the ORT corpus and the increase in students who recognized each word correctly in the post-test. The greatest improvement is the word 'mess' (+6 improvement) in frequency band 13. Amongst the eight words (mess, lorry, thorns, vet, seals, harness, hutch, hay) which no students answered correctly in the pre-test, 'mess' showed the greatest improvement and had the highest frequency. The correlation between the improvement in meaning recognition of these eight words and word frequency was 0.88. As for the six words (troll, pond, pirate, wallpaper, crane, flour) which were each answered correctly by only one student in the pre-test, the correlation between their recognition improvement and frequency was 0.50. Although the sample number is small, the data suggests that a more frequent word may have a better recognition improvement.

Table 4. Improvement of word recognition of each word

Frequency	Target word							
14	balloon							
	0 (9)							
13	mess	elephant	duck					
	+6 (0)	0 (9)	-1 (9)					
12	girl							
	0 (9)							
11	rope	cream	band	hose	scarf			
	+1 (8)	+4 (5)	+2 (5)	0 (3)	+2 (7)			
10	things	hole	fight	troll	lorry			
	+1 (8)	0 (9)	+1 (8)	+2 (1)	+4 (0)			
9	hat	knight	mud					
	0 (9)	+2 (4)	+3 (3)					
8	pond	thorns	pirate	wallpaper				
	+3 (1)	+1 (0)	+5 (1)	+4 (1)				
7	bath	crane						
	+1 (8)	+5 (1)						
6	tap	volcano						
	+3 (2)	+1 (3)						
5	head	sandcastle	ladder					
	0 (9)	+3 (5)	+1 (3)					
4								
3	garden							
	+1 (8)							
2	vet	tail	seals	giraffes	harness			
	+1 (0)	+1 (6)	+1 (0)	+2 (2)	0 (0)			
1	dream	hens						
	0 (9)	+1 (2)						
0	telescope	rubbish	flour	twins	hill	hutch	fist	hay
	-1 (5)	0 (3)	+1 (1)	+1 (6)	+1 (4)	+1 (0)	+1 (3)	0 (0)

\*The figure below each word shows the increase in the number of students who recognized the word correctly. The figure in (brackets) shows the number of students who recognized the word correctly in the pre-test.

The analysis of the relationship between word recognition *improvement* and word frequency, however, cannot be extended much further in this study for the following reason. The hindrance is that as shown in Table 4, many of high frequent words were already answered correctly by many students in the pre-test, giving little chance for improvement in the post-test. For instance, the word 'mess' was the only word in the top frequency bands of 12-14 that did not receive a unanimously correct answer in the pre-test. It is unclear whether the recognition of 'mess' improved the most because it was a high frequency word, or whether some other factors were involved. In addition, several words in the other bands also received a unanimously correct answer in the pre-test, making it difficult to observe any improvement. In future study, sufficient numbers of target words that are previously unknown to participants to represent each frequency band must be provided in order to allow improvement in the post-test.

For the reasons mentioned above, not enough data was collected to discuss the differences between false beginners and non-false beginners in terms of correlation of word recognition improvement and word frequency. One observation made other than the correlation is that where six out of nine students seemed to have learned the word 'mess' between the pre- and the post-test, the three students who did not learn it were all false beginners. This may suggest that certain false beginners have more difficulty in learning a word that other group members found easy. Another point is that two students from the false beginner group and two students from the low-intermediate group made one or two incorrect answers on the words they previously answered correctly in the pre-test. This 'negative' improvement on word recognition was not found in the more advanced group, which may be due to more stable knowledge of vocabulary among the more advanced students.

Although the correlation between *improvement* of word recognition and word frequency was not thoroughly analyzed, the correlation between correct answer rate (the ratio of students who answered a word correctly) and word frequency can be seen in Table 5. The correlation which was 0.35 in the pre-test, improved to 0.53 in the post-test. This may imply that frequent word occurrence may play a part in the enhancement of correct answer rate. The correlation, however, is not strong enough to make any definite conclusion.

**Table 5. Correlation between correct answer rate and word frequency**

<b>Pre-test</b>		<b>Post-test</b>	
Correct answer rate with Word frequency	Correlation 0.35	Correct answer rate with Word frequency	Correlation 0.53

More detailed examination of the correlation between correct answer rate and word frequency

in terms of proficiency levels can be seen in Table 6. The table shows that the correlations in the false beginner and low-intermediate group are significantly stronger than the one in the more advanced group in the pre-test. This discrepancy can be explained by Table 7, which shows the average correct answers per student in the three different frequency bands. In the top frequency band (10-14) of the pre-test, false beginner, low-intermediate and more advanced group averages show little difference. In the middle (5-9) or bottom band (0-4), however, the more advanced group scores much higher than (and in some cases as twice as much as) the other groups. Where the false beginner and the low-intermediate group consistently scores less as the frequency band goes down, the more advanced group's score did not drop consistently nor greatly as the frequency band decreases. In other words, false beginner and the low intermediate groups' word recognition performance seems to be more closely related to word frequency than the more advanced group. In the post-test, the pattern in which the false beginner and low-intermediate groups scored considerably lower in the lower frequency bands whereas the more advanced group scored only a little less than the top frequency band was repeated. This time, however, the more advanced group consistently dropped its scores like the other two groups, giving a relatively stronger correlation than before as seen in Table 6. These data indicate that all three groups increased their correlation between correct answer rate and word frequency. However, a clear relationship between these variables was not found in the study.

Table 6. Correlation between correct answer rate and word frequency according to proficiency groups

Proficiency group	Pre-test			Post-test		
	False beginner	Low-intermediate	More advanced	False beginner	Low-intermediate	More advanced
Correlation between correct answer rate and word frequency	0.43	0.45	0.07	0.53	0.52	0.32

Table 7. Average correct answers per student

Frequency band	Pre-test			Post-test		
	False beginner	Low-intermediate	More advanced	False beginner	Low-intermediate	More advanced
10-14	10	11	9.3	11.6	12.5	13.3
5-9	4.5	5.5	7.0	8.0	5.7	11.0
0-4	4.0	4.5	8.0	4.8	5.0	10.3

### 3.2 Monitoring vocabulary learned during ER and testing the retention rates of the learned words immediately after ER

#### 3.2.1 Words learned during ER

Table 8 shows the number of the words students recorded during ER. The students either guessed meanings or used a dictionary to find out the meanings of unknown words. The use of dictionaries was allowed in ER, but whether to use a dictionary or not, or when to use it, was entirely decided by each individual student. F1, F2 and F4 did not use dictionaries

at all. LI2 and A2 attempted to guess almost all of the recorded unknown words and then used dictionaries on every word to check whether their guesses were correct. LI1 also attempted to guess all the recorded unknown words, but did not use a dictionary on every word. On the other hand, A1 used a dictionary significantly more often than guessing. F4 did not record any words during ER; nor was he able to recall any word he had encountered during ER. Although F4 did not show any evidence of learning new words through ER, the other false beginners F1, F2 and F3 demonstrated that they were able to guess some of the words correctly during ER.

**Table 8. Number of words students recorded during ER**

		F1	F2	F3	F4	LI1	LI2	A1	A2	A3
Total number of words students recorded		9	11	13	0	22	18	122	61	47
Guessing meanings	Correct guessing	8	11	8	0	21	14	3	36	19
	Incorrect guessing	1	0	1	0	0	4	4	20	7
	Total guessing attempts	9	11	9	0	21	18	7	56	26
Referring to a dictionary		0	0	4	0	17	18	115	61	21

Table 9 shows the figures in Table 8 according to the average of each proficiency level. Although the styles of vocabulary learning such as the use of dictionaries vary individually, there are clear differences between the three proficiency level groups.

**Table 9. Number of words students recorded during ER according to proficiency levels**

		F1	F2	F3	F4	LI1	LI2	A1	A2	A3
Total number of words students recorded		8.3				20.0		76.7		
Guessing meanings	Correct guessing	6.8				17.5		19.3		
	Incorrect guessing	0.5				2.0		10.3		
	Total guessing attempts	7.3				19.5		29.6		
Referring to a dictionary		1.0				17.5		65.7		

In all aspects (recording the words during ER, guessing unknown words correctly and incorrectly, referring to dictionaries), more advanced students showed more attempts than low intermediate students, and low intermediate students showed more attempts than false beginners. The clearest discrepancy can be seen in the use of dictionaries, where more advanced students showed significantly more attempts than other groups. In the false beginner group, three of the four students did not use dictionaries at all. This difference may be explained if we suppose that the more advanced students already had a larger vocabulary, so that the words they wanted to refer to in dictionaries during ER were easily identifiable and limited to a certain number; whereas false beginners encountered a larger number of unknown words, which may have made it difficult to refer to dictionaries for

every unknown word. When flooded with unknown words, false beginners may have felt discouraged to use dictionaries altogether. It is possible to suppose that using dictionaries distracted more advanced students very little while false beginners might have felt that the use of dictionaries could be a great distraction. The data may also reflect students' study skills and habits, in that more advanced students may have been more used to using dictionaries, with false beginners not being so accustomed to dictionary use.

### 3.2.2 Retention rate immediately after ER

The words recorded by students were made into a tailored meaning recognition (translation) test to test how well they remembered the words immediately after completing the ER program. The results can be seen in Table 10. With regard to the words students correctly guessed during ER, the average retention rate of the low intermediate group was higher than that of the false beginner group, and the average retention rate of the more advanced group was higher than that of the low intermediate group. It must be noted that F2's retention rate on correctly guessed words was the second highest among all of the participants, after A3. With the exception of the case of F2, however, false beginner students' retention rate was the lowest of the three groups, with F1 not recalling any of the words and F4 not having any words to be tested in the first place (as he did not do any guessing during ER). A1's retention rate was considerably lower than the other two more advanced students. This may be due to the difference in her study habits. A1 only recorded three guessed words during ER but she recorded 115 words with which she used a dictionary. Her study habits may have included a focus on dictionary use when encountering unknown words, rather than on guessing.

Table 10. Retention rate immediately after ER

	F1	F2	F3	F4	L11	L12	A1	A2	A3
Number of words students guessed correctly	8	11	8	0	21	14	3	36	19
Correct answers	0	4	1	-	4	2	0	10	8
Retention rate (%)	0.0	36.4	12.5	-	19.0	14.3	0.0	27.8	42.1
Average retention rate (%)	12.2				16.7		23.3		
Number of words students referred to a dictionary	0	0	4	0	17	18	115	61	21
Correct answers	-	-	0	-	4	2	20	15	7
Retention rate (%)	-	-	0.0	-	23.5	11.1	17.4	24.6	33.3
Average retention rate (%)	0.0				17.3		25.1		
Total number of words students recorded	9	11	13	0	22	18	122	61	47
Correct answers	0	4	2	0	5	2	20	15	17
Retention rate (%)	0.0	36.4	15.4	0.0	22.7	11.1	16.4	24.6	36.2
Average retention rate (%)	13.0				16.9		25.7		

The difference in retention rate between the false beginner group and the other two groups became even clearer with dictionary referred words. The false beginner group's retention rate is zero as three of the four students did not refer to a dictionary during ER, and F3 did not recall any of the four words for which he had used a dictionary. On the other hand, the other two groups' average retention rates were 17.3% for low intermediate and 25.1% for more advanced students. Although LI1's retention rate was higher than A1 and was almost the same level as A2, the actual number of the words LI1 recalled correctly was lower than A1 and A2, as these two students recorded considerably more words during ER than LI1. In short, despite some overlapping retention rates between the low intermediate group and the more advanced group, all of the students in the more advanced group recalled more words than those in the low intermediate group, whereas no students in the false beginner group recalled any words.

Similar analysis can be made for the retention rate of all the words recorded (correctly guessed words, incorrectly guessed words and dictionary referred words). The highest average retention rate is that of the more advanced group and the lowest that of the false beginner group. Although some individual retention rates do not follow this trend (F2's retention rate, for example, is the highest of all participants, and LI1's retention is higher than that of A1), the raw numbers of words that students recalled were highest in the more advanced group. In short, some of the false beginner group students and low intermediate group students had retention rates almost the same as, or higher than, those of more advanced students, but since more advanced students were able to record more words during ER, they remembered more words.

As shown in Table 11, the total number of words students recorded during ER was strongly correlated (0.85) with the amount of reading they were engaged in during the ER program. In other words, the more advanced students read significantly more, recorded more words and remembered more words than students in other groups while false beginner students read less during ER, recorded fewer words (hardly using dictionaries) and remembered fewer words.

**Table 11. Correlation between total recorded words and amount of reading during ER**

	F1	F2	F3	F4	LI1	LI2	A1	A2	A3
Total recorded words	9	11	13	0	22	18	122	61	47
Amount of reading during ER	30,426	33,311	30,781	25,139	34,073	34,075	186,934	198,844	175,843
Correlation between total recorded words and amount of reading during ER	0.85								

## 4 Conclusion

This study saw some possibility of a close relationship between the improvement of word recognition and word occurrence frequency through ER of the ORT picture books. One finding was that a high frequency word, 'mess', in the test gained the greatest improvement in word recognition. Another was that the study saw a correlation of 0.88 between these two factors for some target words. Clear evidence to show the effectiveness of ER for EFL false beginners, however, was not found. Even though the word 'mess' was learned by the greatest number of participant students, this group did not include three out of the four false beginners. Furthermore, even though the average post-ER improvements in word recognition of the three proficiency groups were nearly the same, the lowest improvement in performances was seen in the false beginner group. This study also found that most of the participant students were capable of guessing unknown words or of spontaneously attempting to find out the meanings of unknown words by referring to dictionaries during ER, although the frequencies of these attempts and the accuracy of word guessing varied depending on each student. The data showed that more advanced students recorded more guessed and dictionary referred words during ER. The recall test on these recorded words during ER showed that it was more advanced students who successfully remembered the most words among the three groups and it was the false beginner group students who remembered the fewest words. This study could not explore the factors that cause these different outcomes between the more advanced group and the false beginners. The significantly greater amount of ER reading by the more advanced group, however, seems to play an important role in producing a better result in the recall test as it was the number of total words that students recorded during ER, not the retention rate, that made the difference in the recall test, and the number of total recorded words was strongly correlated with the amount of ER reading. This may suggest that, while the effectiveness of ER in terms of vocabulary development itself must remain the subject of future study, its effectiveness particularly for false beginners, whose reading amount seemed to be hindered by their low English proficiency, could be more limited than for other, more proficient learners, even when using simple picture books such as those of the ORT series.

## References

- Brown, R., Waring, R. and Donkaewbua, S. (2008). Incidental vocabulary acquisition from reading, reading-while-listening, and listening to stories. *Reading in a Foreign Language*, 20 (2), 136-63.
- Kihara, N. (2008). Eigo tadoku reading strategy ni kan suru pilot chousa [Reading strategies in English extensive reading: a pilot study]. *The Japan Association for Developmental Education, 4th National Conference*, Yokohama, 11 August 2008.

Kihara, N. (2009a). Remedial kyouiku toshite no eigo tadoku shidou: chuukan houkoku [Extensive reading as a tool for remedial EFL education: an interim report]. *The Japan Association of College English Teachers, Kyushu-Okinawa Chapter, 23rd Annual Conference*, Okinawa, 20 June 2009.

Kihara, N. (2009b). Extensive reading of picture books and vocabulary development in EFL false beginners. *Pan-Korea English Teachers Association, 2009 International PKETA Conference*, Busan, 10 October 2009.

Kweon, S. and Kim, H. (2008). Beyond raw frequency: Incidental vocabulary acquisition in extensive reading. *Reading in a Foreign Language*, 20 (2), 191-215.