How In-Class Extensive Reading Affects CEFR A2 Level Japanese EFL Learners' Writing: Writing Fluency and Formulaic Sequence Use

Hiroko NISHIMURA

要旨

大学１年生の日本人の初級英語学習者を対象に 20 分の授業内多読が英文ライティングに与える影響を検証した。その結果、産出語数の有意な増加が示された。定型表現は、書き・話し言葉に関係なく産出される傾向が示唆され、事後テストの方が話し言葉に適した表現や句動詞の産出が増える兆候が示された。
キーワード：授業内多読、定型表現、非明示的学習、ライティング

1. Introduction

Fostering communicative competence in English has been identified as a key objective in recent English education in Japan. In this context, knowledge of formulaic sequences (FS; other phrases, such as formulaic expressions, may also be used depending on the context) is considered to play a crucial role (Kadota, 2012). Although the significance of FSs in English language use has been extensively investigated, the research on FS acquisition remains relatively scarce (e.g., Siyanova & Martinez, 2015). Regarding language acquisition, it is widely acknowledged that sufficient comprehensible input is essential (Krashen, 1985), and the same holds for FS acquisition. However, due to the English as a foreign language (EFL) context in Japan, English exposure is significantly limited for most Japanese learners of English. Under such circumstances,

extensive reading (ER) has increasingly attracted attention as a pedagogical approach to providing learners with large amounts of English input. This study investigated the impact of ER on writing development and FS production, to gain insights into a productive skill that has been regarded as Japanese EFL learners' challenge among the four language skills (The Ministry of Education, Culture, Sports, Science and Technology, 2018), targeting Japanese EFL learners with a low proficiency level.

2. Literature Review

2.1 Formulaic Sequences

FSs are frequently used word sequences hypothesized to be stored, retrieved, and processed (without grammatical operations) as single lexicalized units in language use (Wray, 2002). Based on this hypothesis, FS knowledge contributes to the reduction of cognitive load and the enhancement of fluency, accuracy, and communication in language processing. Pawley and Syder were pioneering researchers who highlighted the significance of FSs. They asserted that native speakers of English possess at least hundreds of thousands of FSs and use them appropriately according to the context (Pawley & Syder, 1983). In other words, FSs constitute a substantial proportion of English in use.

Given these factors, FS-knowledge acquisition is essential for enhancing English proficiency. However, it has been argued that FS acquisition is challenging for EFL learners. One reason is that learners' exposure to contextualized FSs is relatively limited, whereas native speakers of English acquire them through repeated use and processing in meaningful communicative contexts from an early age. Furthermore, FSs are characterized by semantic opacity (Moon, 1998), which makes it

difficult to deduce their meanings from their component words. Moreover, since FSs typically comprise common words, such as phrasal verbs, regarded as the most problematic for teaching and learning English (Sinclair, 2004) —consist of monosyllabic verbs and adverbial or prepositional particles, EFL learners frequently fail to notice FSs even when they encounter them (e.g., take place).

From an educational perspective, there are two useful lists: the PHRASal Expressions (PHRASE) List (Martinez & Schmitt, 2012) and the PHrasal VErb Pedagogical (PHaVE) List (Garnier & Schmitt, 2015a, 2015b). FSs that exhibit high frequency and low semantic transparency are important in English language instruction (and learning). Reflecting on this view, Martinez and Schmitt (2012) developed the PHRASE List, which contains 505 FSs based on an analysis of the British National Corpus. Furthermore, the list provides information on the typical contexts in which each FS is used, namely, in spoken or written language, as well as further information on its frequency level, categorized as most common, less common, infrequent, and rare or non-existent. Although such contextual knowledge is crucial for appropriate FS usage, mastering it is an even greater challenge for learners in EFL settings than FS acquisition. Phrasal verbs are the most common (Biber et al., 1999), frequently used, and characterized by high polysemy, many of which are infrequent and peripheral in usage (Garnier & Schmitt, 2015a) . Taking this into account, Garnier and Schmitt (2015a, 2015b) constructed the PHaVE List (of 150 phrasal verbs) derived from the Corpus of Contemporary American English, prioritizing high-frequency phrasal verbs and their most representative meanings for instructional use. From a pedagogical perspective, this study focused on the FSs included in these two lists. Specifically, the production of FSs in the writing of Japanese learners with lower English proficiency levels was

examined by targeting the FSs in the two lists.

2.2 Extensive Reading

In reading, readers are assumed to engage in in several cognitive processes during reading (Kadota et al., 2021). First, they perceive orthographic information (i.e., written text) and mentally convert it into phonological representations (i.e., phonological decoding and lower-level processing). Next, readers access their mental lexicons to retrieve relevant information (of the processed text), such as semantic, syntactic, and related linguistic information, thereby achieving reading comprehension. In phonological decoding, connecting letters and sounds is a prerequisite. Decoding in native languages is automated and requires minimal cognitive resources. By contrast, for a target language, this process often requires considerable cognitive resources. The promotion of phonological decoding automatization requires extensive practice and repeated processing, which gradually reduces the cognitive resources associated with the transformation of orthographic inputs into phonological representations. Accordingly, as phonological decoding in English becomes more automatized, readers can effectively allocate more cognitive resources to other processes such as comprehension (e.g., Just & Carpenter, 1992). Additionally, this enables higher-level processing, such as the expansion of information processing units and the development of chunking skills (Kadota et al., 2021; cf. Nation & Waring, 2019), which leads to the acquisition of new expressions, including FSs. Even in the EFL context, ER can provide plenty of input (Suzuki & Kadota, 2018), serving as a potential approach for providing learners with ample opportunities to engage in phonological decoding.

ER literally refers to reading a large amount of material whose English is easily comprehensible to individual learners (Nation & Waring, 2019).

ER is grounded in Krashen's Input Hypothesis (Krashen, 1985), which emphasizes the importance of comprehensible input for language acquisition. ER materials should be at a level where learners become sufficiently absorbed in the content to forget that they are reading English (Krashen & Mason, 2020). That is, ER books for the early stages were picture books of short, simple English sentences, accompanied by expressive illustrations depicting everyday conversations. By processing limited written input together with visual information, learners can enjoy stories without having to translate them into Japanese. Accordingly, the primary aim of ER is pleasure reading (Day & Bamford, 1998). Studies have reported the effectiveness of ER for English language learners (Takase, 2010), for example, research targeting Japanese EFL learners, improvements in reading speed (Fujita & Noro, 2009), reading ability (Yamashita, 2008), and TOEIC scores (Nishizawa et al., 2006). The effects of ER on writing skills have also been reported (Hafiz & Tudor, 1989). Further, Watanabe and Oba (2018) observed the effects of a four-month in-class ER program on Japanese high school students' English writing skills. The findings suggest that ER facilitates the transfer of writing skills at the word and sentence levels. Overall, these studies imply that ER may implicitly facilitate improvements in learners' English linguistic abilities.

Further, given the ubiquity of FSs, learners are likely to encounter them in authentic and meaningful contexts through ER. Langacker (2008) emphasizes the importance of frequent encounters with FSs in meaningful contexts for FS acquisition. ER offers a valuable means of promoting English learning in EFL settings by enabling learners to access contextually embedded FSs. Based on the foregoing discussion, ER is expected to facilitate FS acquisition; however, few studies have focused specifically on the effect of ER on changes in processing fluency

or the use of FSs (cf. Nation & Waring, 2019) in the writing of Japanese university students with low English proficiency levels. Thus, the present study aimed to examine the impact of ER on the writing performance of low-proficiency English learners (first-year university students). To this end, in-class ER was adopted with the expectation that peer interaction would play a facilitative role, with particular attention to their written output and the production of FSs.

2.3 Research Questions

This study investigates how in-class ER affects beginner-level English learners' writing. The research questions are as follows:

RQ1 (Quantitative analysis) : How does in-class ER affect the writing fluency of beginner-level English learners?

RQ2 (Exploratory lexical analysis) : What kind of FSs are produced in beginner-level English learners' writing after ER?

3. Method

3.1 Participants

This study involved 56 Japanese EFL learners who were first-year university students, none of whom were majoring in English. The author fully explained the research to the students, and all signed written consent forms. After excluding inappropriate students for this study, the total number of students in the analysis was 53. Fifty-two participants' TOEIC Bridge's mean value (M) at the time of university entry was 65.81 (standard deviations: SD = 2.17) , whose converted TOEIC score was 365.29 (SD = 15.95) ; corresponding to the A2 band of the Common European Framework of Reference for Languages scale. All the students were new to ER.

3.2 Procedure

This study was conducted as part of an English course for first-year university students taught by the author (spring and fall semesters; 30 classes, 90 minutes each). Approximately one hour of each class session was devoted to textbook-based instruction that emphasized test-taking strategies and linguistic features, and approximately 30 minutes of class time was allocated to this study's component. Apart from the pre- and posttests, no writing instructions or activities were provided throughout the course.

3.2.1 Pretest and Posttest

Pre- and posttests (unannounced) were conducted in the first and final classes of the spring and fall semesters, respectively. The author distributed a blank A4-sized sheet of paper to each student and asked them to complete a 20-minute writing task. For the pre- and posttests, the prompts were as follows: "Reflect on your high school life/first year at university and describe three memorable experiences freely in English," respectively. To avoid potential misunderstanding, instructions were provided in Japanese. Students were given 3 minutes to prepare after receiving a writing prompt before tackling the task. To avoid the influence of predictive text and spell-check features in the software, the writing task was completed by hand under supervised conditions; no use of smartphones or dictionaries was permitted from the time the prompt was given until the written papers were collected. During the writing task, the web timer was projected onto the classroom screen to inform the students of the remaining time.

After the pretest, the author oriented the students to ER, covering key aspects such as the three fundamental principles of ER (not using a dictionary while ER, skipping the words that you do not understand,

abandoning books that are not enjoyable in favor of more engaging ones）（Sakai, 2002), Sustained Silent Reading（Day & Bamford, 1998), reading for pleasure, book choice, how to fill in the reading record sheet, etc.

3.2.2　Treatment

Starting from the class after the pretest, a 20-minute ER session was conducted during each class, totaling 28 sessions. For ER, the Yomiyasusa Level（YL; Furukawa, 2005）was adopted as the criterion for selecting appropriate ER materials. The YL is a readability scale specifically designed for Japanese learners of English. This is a unique grading system developed by the Japan Extensive Reading Association to support ER practices. Considering their English proficiency levels, the students began reading from YL 0.1. The students were instructed to gradually advance through the YL, systematically following a guide created by the author. For example, the recommended number of books for each YL was as follows: 20 books at YL 0.1, 30 books each at 0.2 and 0.3, 40 books each at YL 0.4, 0.5, and 50 at YL 0.6.

In every class, the author brought a cart containing a sufficient number of books at the appropriate YLs and arranged them at the back of the classroom. To meet the students' varied preferences, a wide selection of readers was made available, such as Oxford Reading Tree（ORT), Building Blocks Library, I Can Read, Foundation Reading Library. Among them, students were primarily encouraged to read the ORT series with a minimum of 12 books per YL（two sets of 6 books each). During the in-class ER, students freely selected books of their choice. Upon finishing a book, they filled out an individual reading record sheet noting details, such as the total number of books read, book title, YL, series name, word count, cumulative word count, and brief comments.

Reading record sheets were collected at the end of each class, along with their class-reflection sheets (on the reflection sheet, students wrote comments on each class's content).

Students were explicitly informed that the primary goal was to enjoy reading, only minimal comments would be required, and their reading volumes would not be reflected in their course grades. Furthermore, owing to the lack of institutional support for ER resources, students were not encouraged to engage in ER outside of class.

3.3 Analysis

The students' handwritten writing was manually transcribed into electronic text for analysis by the author (using Microsoft Excel); during transcription, necessary modifications were made to prepare the texts for analysis. The rules for these modifications were as follows (cf. Hirano, 1990) : 1) Contractions were converted to their full forms (e.g., I'm was transcribed as I am), 2) compound expressions that were mistakenly written as single words were counted according to their correct word boundaries (e.g., highschool transcribed as high school), and 3) proper nouns were counted as a single word (e.g., Tokyo Disney Land). Furthermore, considering the emphasis on writing fluency and the fact that spelling errors are typically auto-corrected when using word-processing software, misspellings were accepted as long as the intended meaning was discernible. Moreover, students were instructed to write as much as possible in English and were required to stop strictly at the end of the allotted time. Therefore, in addition to the completed sentences, meaningful phrases were examined for the analysis. Given the wide variation in writing output, all data within $\pm 3.5SD$ of the mean were retained for the analysis, and no data were excluded. Entries from the reading record sheets and comments from the class reflection sheets

were treated as primary data and analyzed without modification.

Regarding RQ1, to examine the overall features of the students' written output, (due to the sufficient sample size) within-subject *t*-tests and Pearson product-moment correlation coefficient were performed. AntConc 4.3.1 (Anthony, 2023) was used to calculate each student's total number of words on each test.

Next, to address RQ2, an exploratory analysis was conducted. First, the expressions used in the students' writing were examined regarding the FSs in the PHRASE List (Martinez & Schmitt, 2012) and PHaVE List (Garnier & Schmitt, 2015a, 2015b). Then, the FSs from the PHRASE List produced in the students' writing were analyzed for their register (spoken vs. written). Further, to examine the FSs that the students were likely to have encountered during ER, texts from the relevant levels of the ORT (i.e., YL 0.1–0.6: Stages 1–5), which students mainly read, were transcribed manually. To identify whether the FSs (i.e., those listed in the PHRASE and PHaVE lists) produced in students' writing appeared in the ORT materials, each target FS was manually searched across digitized texts (with ORT book-level information) using Microsoft Excel.

As the FSs listed in the PHRASE List are semantically opaque, determining whether a given expression in student writing constitutes a target FS requires careful judgment. Therefore, a researcher with substantial experience in English language teaching served as the interrater. Discrepancies were discussed until a 100% agreement was reached.

4. Results

Table 1 presents the descriptive statistics for students' performance (N = 53) across the 28 in-class ER: the total number of books and words read, and YL with their *M*, *SD*, minimum (*Min*), and maximum (*Max*). Regarding the YLs that students read in the final in-class ER session, eight students reached YL 0.6, having read an average of approximately 19 books at that level. As shown in Table 1, most students read YL books of YL 0.5 during the final ER session.

Table 1
Descriptive Statistics for the Number of Books Read, Words Read, and YLs for 28 In-Class ER

	M	*SD*	*Min*	*Max*
Books	157.96	14.09	127	207
Words	22953.08	5235.7	13997	45893
YL	0.52	0.36	0.5	0.6

Note. N = 53.

4.1 Research Question 1 (Quantitative Analysis)

Table 2 presents the descriptive statistics for the pre- and posttest writing tasks (20 min), including the total number of words, sentences, and words per sentence with their *M* and *SD*. Figure 1 shows the graphical representation of the data of word count (the total number of words) in Table 2. The error bars represent 95% confidence intervals.

Regarding RQ 1, to examine writing fluency, paired *t*-tests were conducted. First, the results concerning the word count are reported. A paired *t*-test result demonstrated a significant difference between the total of words that students wrote in the pre- and posttests, $t(52) = -2.28$,

p = .027, r = .30. This result implies that the students wrote more words on the posttest writing task than on the pretest. Furthermore, paired t-tests were performed to examine differences in the total number of sentences written and the number of words per sentence between the pre- and posttests. Neither comparison yielded a statistically significant difference: $t(52)$ = -0.99, p = .326, r = .14; $t(52)$ = -1.55, p = .128, r = .21, respectively.

To further examine changes in writing performance, Pearson product-moment correlation analyses were conducted on the change scores between the pretest and the posttest for (a) total word count and sentence count and (b) total word count and the number of words per sentence; change scores were calculated by subtracting pretest values from posttest values. Descriptive statistics for these three measures are shown in Table 3. The results of the correlation analysis (a) indicated a strong positive correlation between changes in word count and sentence count, suggesting that students who wrote more words tended to produce more sentences: $r(51)$ = .85, p < .001. Meanwhile, correlation analysis (b) yielded a weak negative correlation, indicating that increases in word count were slightly associated with decreases in the average number of words per sentence: $r(51)$ = -.25, p = .068.

Table 2
Descriptive Statistics for Total Number of Words, Sentences, Words per Sentence in Pre- and Posttests

	Word Count		Sentence Count		Words/Sentence	
	M	*SD*	*M*	*SD*	*M*	*SD*
Pretest	102.15	37.76	13.55	6.04	7.74	1.49
Posttest	111.38	42.90	14.21	6.04	8.15	2.06

Note. N = 53.

Figure 1
Total Number of Words Written in Pre- and Posttests

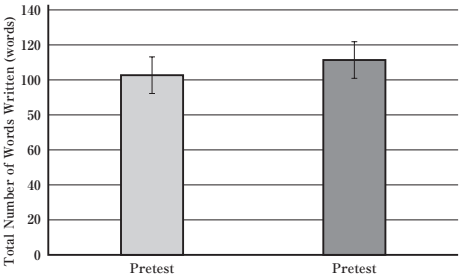

Table 3
Descriptive Statistics for the Changes of the Total Number of Words, Sentences, and Words per Sentence Between the Pretest and the Posttest

	Word Count Change	Sentence Count Change	Words/Sentence Change
M	9.23	0.66	0.40
SD	29.43	4.85	1.90

Note. $N = 53$.

4.2 Research Question 2 (Exploratory Lexical Analysis)

4.2.1 Produced FSs in the PHRASE List

The total number of FS tokens found in the pre- and posttest writing was 44 and 48, respectively (the four students who used a single FS twice, two in the pretest and two in the posttest, were all different individuals; each of the four repeated FSs was unique, with no overlap between the tests). Table 4 summarizes the number of tokens of FSs produced in the pre- and posttests, categorized by the PHRASE List's written and spoken frequency levels (*** = most common or common, ** = less common, * = infrequent). Although the PHRASE List includes a written academic genre, this category was excluded from the current analysis, considering both students' English proficiency levels and the

objectives of the study. As shown in Table 4, a limited number of FSs was produced, most of which were commonly used in either written or spoken languages. Furthermore, the total number of FS tokens did not increase significantly from the pretest to the posttest. Regarding the most commonly written FSs category, the number of tokens remained the same in both tests. There appears to be a slight increase in tokens categorized as the most commonly spoken FSs from the pretest to the posttest.

The types of FSs produced in the pre- and posttests were 30 and 24, respectively (18 of which were common to the two tests). Table 5 presents examples of FSs produced in the pre- and posttests categorized by the PHRASE List of frequency levels for written and spoken English, along with their token frequencies (numbers in parentheses). FSs underlined in the table indicate those that appeared in ORT (YL 0.1–0.5). A possible increase was observed in FSs classified in the PHRASE List as less common in written language but most common in spoken language in the posttest writing.

Table 4
FS Token Counts in Pre- and Posttests by Written and Spoken Frequency

Written × Spoken	***	**	*	Row Total
***	18/19	4/5	4/2	26/26
**	12/18	0/0	1/0	13/18
*	5/4	0/0	0/0	5/4
Column Total	35/41	4/5	5/2	44/48

Note. Each cell shows the number of tokens in the students' pre- and posttest writing (N = 53), respectively (pretest/posttest). The rows represent written frequency categories and the columns represent spoken frequency categories.

Table 5
Examples of Produced FSs by Genre Frequency and Test Phase

PHRASE List Frequency		Pretest Only	Posttest Only	Both (Pretest, Posttest)
written***	spoken***	not only (1) take place (1) instead of (1) on the other hand (1) thanks to (2)	at all (1)	there is/are (5†, 6) come back (1, 1) in front of (2, 2) and so on (1, 3) at first (2, 3) these days (1, 3)
written***	spoken**	N/A	N/A	in addition (1, 2) in particular (2, 1) look forward to (1, 2)
written***	spoken*	such as (1) for instance (1)	N/A	take part in (2, 2)
written**	spoken***	look for (1) think about (1) get to (arrive at) (2†)	going to (FUTURE) (2) at least (1) pick up (1) look like (1) right now (1)	have to (3, 4) of course (2, 1) a little (1, 2) each other (1, 2†) get up (1, 3†)
written**	spoken**	N/A	N/A	N/A
written**	spoken*	as a result (1)	N/A	N/A
written*	spoken***	first of all (2)	N/A	a lot (1, 1) used to (PAST) (1, 1) good at (1, 2)
written*	spoken**	N/A	N/A	N/A
written*	spoken*	N/A	N/A	N/A

Note. N/A = No occurrences observed in students' writing. † The same FS was used by the same student.

Table 6
Examples of Produced PVs With Their Test Phase

Rank	PVs	Pretest	Posttest
2	pick up		1
3	come back	1	1
16	give up		1
23	get up	1	3*
35	wake up		1
115	go around		2*
Total		2	9

Note. Rank = position of each PVs in the PHaVE list; * The same PV was used twice by the same student. Underlined PVs indicate those that appeared in ORT (YL 0.1–0.5).

4.2.2　Produced PVs in the PHaVE List

Table 6 presents examples of PV types and token frequencies produced in the pre- and posttests. In the table, “Rank” denotes the position of each PV in the PHaVE List. Underlined PVs indicate those that appeared in ORT (YL 0.1–0.5). As shown in Table 6, the number of PVs produced is very limited. However, the posttest results may suggest a potential tendency toward an increase. Only two students used the target PVs in the pretest, and seven students did so in the posttest. No student produced the target PVs in either test.

4.2.3　Students’ Reflection Sheets

Although most students initially lacked confidence in English learning and had low motivation, their reflection sheets indicated positive perceptions after experiencing ER for the first time. The most frequent remarks were, “I can now understand short phrases in English directly in English” and “Reading has become much faster and easier because I can process English without translating.” Other common comments were of the type, “I no longer feel resistance to English.” Additionally, some students simply wrote, “It was fun” or “I liked it.”

5.　Discussion and Conclusion

This study examined the effects of in-class ER on the writing of beginner-level Japanese EFL learners. The quantitative results suggest that ER facilitated an increase in the number of words produced, indicating a positive effect on writing fluency. This result aligns with previous research showing ER’s potential to promote writing fluency. Although the number of sentences and words per sentence did not change significantly, the correlation analyses implied that students who

wrote more words tended to produce more sentences. This outcome may reflect the influence of the ER materials they read, which mainly consisted of short, narrative-style sentences tailored for beginners. Therefore, an exploratory analysis was conducted on the FSs listed in the PHRASE List (Martinez & Schmitt, 2012; hereafter FSs in this section) and phrasal verbs in the PHaVE List (Garnier & Schmitt, 2015a, 2015b; hereafter PVs in this section) produced in the students' writing. The results implied that students' FS and PV production remained limited, with most expressions belonging to the high-frequency categories in either the spoken or written registers. From the perspective of the written and spoken registers of FSs, students seemed to produce FSs regardless of whether they were typically associated with written or spoken usage. When comparing the pretest and the posttest, FSs that were produced only in the posttest showed a tendency toward spoken-register FSs, and even with limited data, there were tentative signs of increased use of PVs. Some of these newly produced FSs and PVs were also found in ORT texts at the target level. This may reflect the influence of beginner-level ER books, which largely consist of narratives and thus naturally contain a high proportion of conversational expressions.

The subsequent discussion was framed by two key conditions: the participants were beginner-level EFL learners, and the average number of words read in-class ER amounted to only approximately 25,000. Given these conditions, the findings should be interpreted as early-stage ER outcomes. The improvement in writing fluency observed from the pretest to the posttest, together with a strong positive correlation with sentence production (number), FS, and PV use, can be interpreted in relation to the following four possible contributing factors. The first possible factor is the more efficient use of cognitive resources through

the use of formulaic expressions. As reviewed in Section 2.2, Watanabe and Oba (2018) suggest that Japanese EFL learners may transfer forms encountered in ER to their own writing, particularly at the word and sentence levels. Additionally, from a cognitive processing perspective, ER can serve as a training tool in phonological decoding, and the gradual expansion of processing units may facilitate the internalization of chunks and formulaic expressions. It is therefore possible that the input processing afforded by ER, combined with repeated exposure to diverse formulaic expressions, including targeted FSs and PVs, facilitated learners' ability to draw on and reproduce such prefabricated expressions in their writing. Second, in writing, students could process words and short phrases directly in English. Picture books may have supported implicit learning by linking visual images with written text, enabling students to directly associate certain expressions, including phrasal verbs, in English. Moreover, the ability to process short phrases directly in English may have contributed to faster rereading during writing. This, in turn, could reduce the cognitive load involved in monitoring output. The third possible factor is the influence of students' positive psychological changes toward English, such as reduced anxiety and increased motivation. This may have facilitated greater engagement with writing. Finally, interactions may have also played a role. Shared experiences of in-class ER may have encouraged peer interactions and fostered motivation. While this is not a direct effect of ER, the stronger rapport that developed between the instructor and students by the end of the 30 lessons, compared to the initial pretest session, may have reduced anxiety and facilitated greater ease of writing.

Overall, the present study indicates that even at the early stage of ER may contribute to improved writing fluency and influence the use of certain formulaic expressions. Notably, the production of phrasal verbs

shows initial signs of change. Although these findings are still tentative, they point to the potential and highlight the need for further research to clarify these emerging trends.

These findings lead to pedagogical implications. Given learners' production of FSs and PVs, it seems important for teachers to recognize that many ER books, especially at the beginning level, are narrative in nature and contain a large proportion of spoken expressions. Thus, introducing nonfiction readers according to learners' proficiency levels may be beneficial (cf. Nation & Waring, 2019).

6. Limitations and Further Study

This study has three limitations. The first is the influence of out-of-class factors. Nation and Waring (2019) argue that such factors are unavoidable in ER research and must be accepted despite efforts to control them. Second, FSs not included in the two target lists were excluded from the analysis. Finally, the findings are limited to specific writing topics and participants. Future research should examine how continued ER influences writing and FS acquisition, thereby clarifying ER's contribution to Japanese EFL learners' overall proficiency.

Acknowledgements

This work was supported by JSPS KAKENHI Grant Number JP23K12257. I would also like to thank my students for their participation in this study.

References

Anthony, L (2023). *AntConc* (Version 4.3.1) [Computer software]. https://www.laurenceanthony.net/software

Biber, D., Jonansson, S., Leech, G., Conrad, S., & Finegan, E. (1999). *Longman grammar of spoken and written English*. Longman.

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

Fujita, K., & Noro, T. (2009). The effects of 10-minute extensive reading on the reading speed, comprehension and motivation of Japanese high school EFL learners. *ARELE, 20*, 21–30. https://doi.org/10.20581/arele.20.0_21

Furukawa, A. (Ed.). (2005). *Tadokutosho YL・gosu risuto* [YLs and word counts list of books for extensive reading] (3rd ed.). Japan Extensive Reading Association.

Garnier, M., & Schmitt, N. (2015a). The phave list: A pedagogical list of phrasal verbs and their most frequent meaning senses. *Language Teaching Research, 19*(6), 645–666. https://doi.org/10.1177/1362168814559798

Garnier, M., & Schmitt, N. (2015b). The phave list: A pedagogical list of phrasal verbs and their most frequent meaning senses [PDF]. https://docs.wixstatic.com/ugd/5f2482_ba18c227594d463aae9438e7b065f592.pdf

Hafiz, F. M., & Tudor, I. (1989). Extensive reading and the development of language skills. *ELT Journal, 43*(1), 4–14. https://doi.org/10.1093/elt/43.1.4

Hirano, K. (1990). Gengonoryoku no kyakkantekisihyo no datosei: Nihonjin EFL daigakuseinobaai [The varidity of objective measures of language proficiency: The case of Japanese EFL university students]. *Joetsukyoikudaigakukenkyu kiyo, 9*(2), 65–77.

Just, M. A., & Carpenter, P. A. (1992). A capacity theory of comprehension: Individual differences in working memory. *Psychological Review, 99*(1), 122–149. https://doi.org/10.1037/0033-295x.99.1.122

Kadota, S. (2012). *Shadowing ondoku to eigogakushu no kagaku: inputto kara autoputto e* [Shadowing and oral reading in English learning: From input to output]. Cosmopier.

Kadota, S., Takase, A., & Kawasaki, M. (2021). *Eigo riding no ninchikagaku: mojigakushu to tadoku no koka o saguru* [The cognitive science of English reading: Exploring the effects of word learning and extensive reading]. Kuroshio.

Krashen, S. D. (1985). *The input hypothesis: Issues and implication*. Longman.

Krashen, S., & Mason, B. (2020, May). The optimal input hypothesis: Not all comprehensible input is of equal value. CATESOL Newsletter. http://beniko-mason.net/content/articles/2020-the-optimal-input-hypothesis.pdf

Langacker, R. W. (2008). Cognitive grammar as a basis for language instruction. In P. Robinson & N. C. Ellis (Eds.), *Handbook of cognitive linguistics and second language acquisition* (pp. 66–88). Routledge.

Martinez, R., & Schmitt, N. (2012). A phrasal expressions list. *Applied Linguistics, 33*

(3), 299–320.

Ministry of Education, Culture, Sports, Science and Technology. (2018). *Heisei 29nendo eigoryoku chosakekka koko 3nensei no gaiyo* [Summary of the English proficiency survey results for third-year high school students in the heisei 29 academic year]. https://www.mext.go.jp/a_menu/kokusai/gaikokugo/__icsFiles/afieldfile/2018/04/06/1403470_03_1.pdf

Moon R. (1998). Vocabulary connections: Multi-word items in English. In N. Schmitt & M. McCarthy (Eds.), *Vocabulary: Description, acquisition and pedagogy* (pp. 40–63). Cambridge University Press.

Nation, I. S., & Waring, R. (2019). *Teaching extensive reading in another language*. Routledge.

Nishizawa, H., Yoshioka, T., & Itoh, K. (2006). *Eigotadoku niyoru kogakukeigakusei no eigo unyonoryokukaizen* [Improvement of engineering students' communication skill in English through extensive reading] . *IEEJ Trans. FM, 126*(7), 556–562. https://doi.org/10.1541/ieejfms.126.556

Pawley, A., & Syder, F. H. (1983). Two puzzles for linguistic theory: Nativelike selection and nativelike fluency. In J. C. Richards & R. W. Schmitt (Eds.), *Language and Communication* (pp. 191–226). Longman.

Sakai, K. (2002). *Kaidoku 100 man go! Pepabakku eno michi* [Reading one million words for pleasure: The path to paper backs]. Chikuma Gakugei Bunko.

Sinclair, J. (2004). *Trust the text: Language, corpus and discourse*. Routledge.

Siyanova-Chanturia, A., & Martinez, R. (2015). The idiom principle revisited. *Applied Linguistics, 36*(5). 549–569. https://doi.org/10.1093/applin/amt054

Suzuki, J., & Kadota, S. (2018). *Eigo risuningu shido hando bukku* [Handbook of instruction of English listening] (pp. 335–345). Taishukan.

Takase, A. (2010). *Eigo tadoku・tacho shido manyuaru* [A handbook of English extensive reading and listening]. Taishukan.

Watanabe, M., & Oba, H. (2018). Kyoshitsunaieigotadoku ga nihonjinkokosei no sakubunryokuni ataeru koka [The effects of in-class extensive reading on the writing ability of Japanese high school students]. *Nihonkyokakyoikugakkaishi, 14*(1), 73–84. https://doi.org/10.18993/jcrdajp.41.1_73

Wray, A. (2002). *Formulaic language and the lexicon*. Cambridge University Press.

Yamashita, J. (2008). Extensive reading and development of different aspects of L2 proficiency. *System, 36*(4), 661–672. https://doi.org/10.1016/j.system.2008.04.003