The Relationship between Text Levels and Collocation Levels in English Textbooks

Risako Azemoto

Department of Integrated Sciences for Global Society, Kyushu University, Japan
Email: risako.azemoto@gmail.com (R.A.)

Abstract—The purpose of this study is to elucidate the relationship between the CEFR levels of English textbooks and the levels of collocations. This study investigated the CEFR levels at which collocations first appear, using corpora from CEFR-based textbooks and Japanese high school English textbooks. I found that a significant minority of the collocations in Japanese English textbooks matched their CEFR levels to those of the texts, whereas over half did so in CEFR-based textbooks. However, approximately half of the collocations in Japanese English textbooks and one-third of those in CEFR-based textbooks were introduced to learners who were not proficient enough. The two types of English textbooks exhibited different transitions in collocation levels, with CEFR-based textbooks better aligning with learners' expected proficiency levels when introducing collocations.

Keywords—CEFR, collocations, corpus linguistics, English textbooks

I. INTRODUCTION

It is evident that vocabulary acquisition is essential in learning English; however, learning collocations, which function similarly to single words but consist of multiple words, is also widely recognized as beneficial for second language learners [1, 2]. Mastering collocations can significantly enhance both comprehension and production in a second language [3]. Notably, it has been shown that collocations are challenging to acquire, even for advanced learners [4–6].

Collocations can be learned not only through explicit instruction but also incidentally through exposure in context [7, 8]. Textbooks play a central role in language learning [9], and it is crucial to address the treatment of collocations within these materials. However, research on collocations in English textbooks in Japan remains relatively undeveloped.

Unlike vocabulary studies, comprehensive analyses of collocations in Japanese English textbooks are scarce. Takesue [10] analyzed collocations involving verbs in one series of junior high school English textbooks, considering their types, parts of speech, and frequency. Subsequently, Takesue [11] expanded this analysis to include six series of junior high school English textbooks, examining changes from older to newer editions in four of these series. Azemoto [12] clarified the range, frequency, and coverages of two collocation lists for basic verb collocations found in six series of high school English textbooks.

Although these studies have analyzed collocations in English textbooks with a focus on parts of speech and frequency, research addressing the levels of collocations is absent. Therefore, this study aims to clarify the relationship between the CEFR levels of English textbooks and the CEFR levels of collocations that appear in them.

II. RESEARCH

A. Research Questions

To achieve this aim, this study will address the following research questions: (1) How are collocations used in CEFR-based English textbooks in terms of CEFR levels? (2) How are collocations used in Japanese high school English textbooks in terms of the CEFR levels? (3) What are the differences between CEFR-based English textbooks and Japanese high school English textbooks in terms of the CEFR levels of the collocations used?

B. Data

1) CEFR-based English textbook corpus

This corpus has been independently created by the author and comprises a total of 208,321 words, extracted solely from the reading sections of CEFR-based English textbooks that are widely used across the world by people of all ages, excluding optional lessons. Twenty-two textbooks from four series—Cambridge English Empower [13], Headway [14], Life [15], and NEW Language Leader [16]—were chosen for inclusion in this corpus. CEFR levels are annotated in all of these textbooks by the authors: EMPOWER has textbooks at levels A1, A2, B1, B1+, B2, and C1; Headway has textbooks at levels A1, A1-A2, A2-B1, B1, B2, and C1; LIFE has textbooks at levels A1, A2, B1, B1+, B2, and C1; and NEW Language Leader has textbooks at levels A1-A2, A2+-B1, B1-B1+, B2, and B2+-C1.

To classify these textbooks into CEFR levels, the texts in textbooks that span two CEFR levels were processed by the CVLA [17]. The CVLA is an online application designed to estimate the CEFR-J level of input texts. The estimation method is based on four indicators—ARI (Readability Index), VperSent (average number of verbs per sentence), AvrDiff (average vocabulary level based on the CEFR-J WordList), and BperA (ratio of B-level content words to A-level content words). When users input a text, they receive values for the four indicators and the corresponding CEFR levels, and the average of those four levels is given as the final text level. In this analysis, the final CEFR level was assigned by choosing the one closest to the level calculated by the CVLA from the two CEFR levels indicated in the textbooks. For example, texts in the NEW Language Leader A2+-B1 textbook that were annotated as A2 by the CVLA were classified as A2, those annotated as B1 were classified as B1, but those annotated as B2 were classified as B1 because the set CEFR level of the textbook is A2+-B1, and B2 is out of range.

The reason some texts were not classified as annotated by the CVLA is that the CVLA is possibly not a perfect CEFR-level analyzer, as Azemoto and Uchida [18] pointed out. This study used the CVLA within the textbooks' CEFR

range, resulting in some texts being classified differently. Table 1 shows the number of texts, tokens, and average token per text at each CEFR level.

Table 1. An overview of CEFR-based English textbook data

	A1	**A2**	**B1**	**B2**	**C1**	**Total**
Text	68	94	186	116	97	561
Token	14,325	30,039	81,570	61,301	59,350	246,585
Token /text	210.7	319.6	438.5	528.5	611.9	2,109.1

2) Japanese high school English textbooks corpus

This corpus has been also independently created by the author and consists of a total of 202,052 words, extracted solely from the reading sections of government-approved Japanese high school English textbooks, excluding optional lessons. These textbooks were selected to achieve an adoption rate of approximately 80%, based on the selection results by Tokyo Metropolitan High Schools for the 2016 academic year. Specifically, 12 textbooks were chosen for the first year, 14 for the second year, and 11 for the third year.

To classify the texts in the corpus, the CVLA was used to measure the text levels. In this study, PreA1 level texts were integrated into A1, and C2 level texts were integrated into C1, as shown in Table 2.

Table 2. An overview of Japanese English textbook data

	A1	**A2**	**B1**	**B2**	**C1**	**Total**
Text	49	95	135	86	53	418
Token	12,173	46,791	81,176	60,365	34,153	234,658
Token /text	248.4	492.5	601.3	701.9	644.4	561.4

3) English vocabulary profile

"Collocation" has various definitions, and previous studies have used the term in different ways. This study focused solely on the phrases and phrasal verbs in the English Vocabulary Profile (EVP) [19]. The EVP is a web resource that provides a comprehensive database of vocabulary information for English learners. It offers details on which vocabulary can be used by learners at each CEFR level. The list includes 15,696 vocabulary items. The CEFR level of each vocabulary item was annotated using various criteria. Although the detailed process of annotating CEFR levels is not publicly available, the EVP explains that it is mainly based on evidence from the Cambridge Learner Corpus.

The EVP includes 3,656 phrases and 728 phrasal verbs. EVP selected these phrases with the definition of a phrase as "a lexical chunk of some kind (i.e., a group of two or more words which frequently occur together) where the meaning of the chunk differs in some way from the meaning of each word added together. [19]" This might be a unique criterion as other collocation lists or phrase lists tend to be based only on frequencies or statistical values in a general or specific corpus.

Although the EVP contains phrases and phrasal verbs composed of more than two words, this study focused exclusively on two-word phrases and phrasal verbs. The phrases and phrasal verbs in EVP sometimes include abbreviations such as "sb" (somebody) and "sth" (something). To make the data analyzable for this study, these abbreviations were first removed from the list. Next, phrases containing three or more words were deleted. Finally, the

remaining candidates were manually checked to ensure they were suitable for this study. As a result, 596 phrases and 500 phrasal verbs were included.

Some phrases and phrasal verbs appear in the EVP more than once. This is because when a phrase or phrasal verb has multiple meanings, the EVP annotates a CEFR level for each meaning, resulting in duplicated entries. In this study, duplicated entries were merged, with the lowest CEFR level being annotated.

C. Data Analysis

To address the research questions in this study, it is necessary to analyze the frequencies of the target collocations in the two corpora. Python 3.10.12, specifically the spaCy 3.7.4 with the en_core_web_sm model and Pandas 2.0.3 libraries, was used for the analysis. The collocations that are directly adjacent to each other, as well as those where one or two words are positioned between the two words of the collocation, were searched in the lemmatized texts in the two corpora.

D. Results

Table 3 presents a part of the result, indicating which collocations appear in which text files. For example, as shown in the first row, in the Headway (HW) A2 level text, the collocation "book in" at the B1 level appears once, and the collocation "call in" at the B2 level appears once.

Table 3. A part of the result

Textbook	Text level	Collocation	Collocation level	Count
HW	A2	book in	B1	1
HW	A2	call in	B2	1
HW	A2	carry out	B1	1

1) How are collocations used in CEFR-based English textbooks in terms of CEFR levels?

Table 4. Phrase collocations in CEFR-based English textbooks

		Text level					Total	Level Match	Level Inappro -priate
		A1	A2	B1	B2	C1			
Collocation level	A1	10	1	2	0	0	13	76.9	0.0
	A2	9	21	10	0	1	41	73.2	0.0
	B1	8	26	39	7	3	83	78.3	9.6
	B2	6	30	36	17	6	95	55.8	37.9
	C1	4	9	15	18	7	53	47.2	52.8
	C2	6	5	16	15	9	51	17.6	82.4
Total		43	92	118	57	26	336	57.1	33.9

Table 5. Phrasal verb collocations in CEFR-based English textbooks

		Text level					Total	Level Match	Level Inappro -priate
		A1	A2	B1	B2	C1			
Collocation level	A1	4	0	0	0	0	4	100.0	0.0
	A2	7	7	7	1	1	23	60.9	0.0
	B1	16	17	45	10	5	93	66.7	17.2
	B2	16	24	51	29	21	141	56.7	28.4
	C1	3	2	11	11	7	34	52.9	47.1
	C2	2	4	21	5	13	45	28.9	71.1
Total		48	54	135	56	47	340	56.2	30.6

Tables 4 and 5 show the number of phrase collocations and phrasal verb collocations, respectively, that first appear at different levels within the CEFR-based textbooks. For

example, there are a total of 13 phrase collocations at A1 level, with 10 of those first appearing in A1 level textbooks, 1 of those first appearing in A2 level textbooks, and 2 of those first appearing in B1 level textbooks.

The "Level Match" column indicates the proportion of collocations where the collocation's level matches the level of the text in which it first appears, or where the collocation's level is only one level higher, for example, "base on" (B1) and "believe in" (B1). The reason for including cases where the collocation's level is only one level higher in the "Level Match" is that the EVP levels are primarily assigned based on learner corpora. This means the assigned level indicates the level at which learners are deemed capable of "using" the collocation. Therefore, textbooks might judge that, in addition to introducing collocations of the same level, it is necessary to include collocations from the next level when introducing collocations within reading comprehension sections.

Overall, 192 out of 336 phrase collocations (about 57.1%) and 191 out of 340 phrasal verb collocations (about 56.2%) were judged as Level Match. The Level Match ratio tends to decrease as the collocation level increases. This indicates that the level of the texts in which higher-level collocations first appear may not be appropriate. This trend was observed for both phrasal verb collocations and phrase collocations. I conducted a paired *t*-test using R (version 4.1.0) with the "*t*.test" function from the "stats" package (version 4.1.0), using the default setting of a two-sided alternative hypothesis. It compared the Level Match ratios for phrase collocations and phrasal verb collocations and found no significant difference. In particular, the ratio drops significantly at the C2 level. This is partly because, whereas there are C2 level collocations, there are no C2 level textbooks. Therefore, a complete match between the collocation level and the text level cannot occur, and only cases where the collocation level is one level higher than the text level can be counted as a Level Match.

The "Level Inappropriate" column indicates the proportion of collocations that appear in texts at least two levels below the collocation's designated level, for example, "fresh water" (C2) and "if so" (B2). Collocations deemed "Level Inappropriate" are those considered too difficult for expected learners using the textbook. Specifically, B1 level collocations are considered inappropriate in A1 level texts, B2 level collocations are inappropriate in A1 and A2 level texts, C1 level collocations are inappropriate in A1, A2, and B1 level texts, and C2 level collocations are inappropriate in A1, A2, B1, and B2 level texts. The reason there are no "Level Inappropriate" collocations at the A1 and A2 levels is that it is believed learners are mature enough to learn A1 and A2 collocations regardless of whether they appear in texts from A1 to C1 levels.

Overall, 114 out of 336 phrase collocations, or approximately 33.9%, and 104 out of 340 phrasal verb collocations, or about 30.6%, first appeared at an inappropriate text level. This indicates that at least 30% of the collocations appeared for the first time at an inappropriate level. A major trend is that the proportion of Level Inappropriate collocations increases as the collocation level rises. One contributing factor is that as the collocation level rises, the number of text levels deemed inappropriate increases as well, but even considering this, it becomes clear that the higher the collocation level, the more difficult it is to introduce collocations at an appropriate level. A paired *t*-test

conducted on the proportion of "Level Inappropriate" collocations for phrase collocations and phrasal verb collocations found no significant difference. This indicates that both phrase collocations and phrasal verb collocations follow similar trends.

2) How are collocations used in Japanese high school English textbooks in terms of CEFR levels?

Table 6. Phrase collocations in Japanese English textbooks

		text level					Total	Level Match	Level Inappro-priate
		A1	A2	B1	B2	C1			
collocation level	A1	7	5	0	0	0	12	58.3	0.0
	A2	22	14	6	2	1	45	80.0	0.0
	B1	29	39	17	3	1	89	62.9	32.6
	B2	23	33	21	18	3	98	39.8	57.1
	C1	15	18	10	11	4	58	25.9	74.1
	C2	12	20	11	8	7	58	12.1	87.9
Total		108	129	65	42	16	360	44.4	49.7

Table 7. Phrasal verb collocations in Japanese English textbooks

		text level					Total	Level Match	Level Inappro-priate
		A1	A2	B1	B2	C1			
collocation level	A1	3	1	0	0	0	4	75.0	0.0
	A2	6	12	4	0	1	23	78.3	0.0
	B1	14	41	29	7	3	94	74.5	14.9
	B2	28	44	30	17	10	129	36.4	55.8
	C1	3	10	8	6	0	27	22.2	77.8
	C2	7	12	19	11	4	53	7.5	92.5
Total		61	120	90	41	18	330	44.8	47.3

Tables 6 and 7 show the number of phrase collocations and phrasal verb collocations, respectively, first appearing at each level within the Japanese high school English textbooks. Overall, 160 out of 360 phrase collocations (approximately 44.4%) matched the textbook level at which they first appeared. For phrasal verb collocations, 148 out of 330 collocations (approximately 44.8%) matched the text level. This indicates that around 44% of the collocations appeared at the appropriate text levels when they first appeared, for example, "at least" (A2) and "due to" (B1).

When examining each level, both phrase collocations and phrasal verb collocations had the highest level match ratio at the A2 level. This ratio drops sharply at the B2 level. The general trend shows that as the collocation level increases, the Level Match ratio decreases. This suggests that it becomes more challenging to introduce high-level collocations at the appropriate text level. The sharp decrease in the Level Match ratio beyond the B2 level is notable. According to the Japanese government [20], the instructional content of high school English textbooks ranges from A1 to B1 levels, which might explain why collocations at the B2 level and above are less likely to be introduced at the appropriate level. A paired *t*-test conducted on the Level Match ratios for phrase collocations and phrasal verb collocations showed no significant difference, indicating that both follow similar trends.

179 out of 360 phrase collocations (approximately 49.7%) and 156 out of 330 phrasal verb collocations (approximately 47.3%) were regarded as Level Inappropriate collocations, for example, "even if" (B2) and "eye contact" (C2). This means that about half of the collocations first appeared in

texts read by learners who had not yet reached the level to properly learn those collocations. Similar to the analysis of CEFR-based textbooks, the proportion of Level Inappropriate collocations increases as the collocation level rises. Although this is partly due to the increase in the number of text levels deemed inappropriate as the collocation level rises, it also shows that Japanese English textbooks are not introducing collocations to learners at the correct level as the difficulty of the collocations increases. A paired *t*-test conducted on the proportions of Level Inappropriate collocations for phrase collocations and phrasal verb collocations found no significant difference. This indicates that both phrase collocations and phrasal verb collocations follow similar trends.

3) What are the differences between CEFR-based English textbooks and Japanese high school English textbooks in terms of the CEFR levels of the collocations used?

Table 8 shows the results of comparing the Level Match ratio between CEFR-based textbooks and Japanese English textbooks. Overall, CEFR-based textbooks have approximately 12.7% higher in phrase collocation and 11.4% higher in phrasal verb collocation than those in Japanese English textbooks. When looking at each level, CEFR-based textbooks are higher in all levels except for A2 level phrase collocations and A2 and B1 level phrasal verb collocations. Overall, as the level increases, the Level Match Ratio tends to decrease, which is consistent across both textbook types. Paired *t*-tests were conducted for phrase collocations and phrasal verb collocations, showing a significant difference in phrase collocations ($p = 0.042$) but not in phrasal verb collocations ($p = 0.194$).

Table 8. A comparison of level match ratio

Collocation level	Phrase collocation		Phrasal verb collocation	
	CEFR-based	Japanese high school	CEFR-based	Japanese high school
A1	76.9	58.3	100.0	75.0
A2	73.2	80.0	60.9	78.3
B1	78.3	62.0	66.7	74.5
B2	55.8	39.8	56.7	36.4
C1	47.2	25.9	52.9	22.2
C2	17.6	12.1	28.9	7.5
Total	57.1	44.4	56.2	44.8

Table 9 compares the Level Inappropriate ratio between CEFR-based textbooks and Japanese English textbooks. Overall, there is a difference of approximately 15.8% in phrase collocations and 16.7% in phrasal verb collocations, with CEFR-based textbooks being lower in both. When looking at each level, CEFR-based textbooks are lower in all values except for the B1 level phrasal verb collocations. The trend of increasing values with higher collocation levels is consistent in both textbook types. Paired *t*-tests were conducted for phrase collocations and phrasal verb collocations, showing a significant difference in phrase collocations ($p = 0.023$) but not in phrasal verb collocations ($p = 0.081$).

In both phrase collocations and phrasal verb collocations, it is clear that CEFR-based textbooks generally have a higher level match ratio and a lower Level Inappropriate ratio. Therefore, compared to Japanese high school English

textbooks, it can be seen that CEFR-based textbooks consider the assumed proficiency level of the learners reading the texts when introducing collocations. The significant differences in phrase collocations, as opposed to the non-significant differences in phrasal verb collocations, indicate that this tendency is particularly pronounced in phrase collocations.

Table 9. A comparison of level inappropriate ratio

Collocation level	Phrase collocation		Phrasal verb collocation	
	CEFR-based	Japanese high school	CEFR-based	Japanese high school
A1	0.0	0.0	0.0	0.0
A2	0.0	0.0	0.0	0.0
B1	9.6	32.6	17.2	14.9
B2	37.9	57.1	28.4	55.8
C1	52.8	74.1	47.1	77.8
C2	82.4	87.9	71.1	92.5
Total	33.9	49.7	30.6	47.3

III. CONCLUSION

In this paper, a collocation analysis was conducted on CEFR-based textbooks and Japanese high school English textbooks. The results showed that from the CEFR-based textbooks, approximately 57.1% of the phrase collocations and 56.2% of the phrasal verb collocations were first introduced at a level that matched their collocation level. In contrast, about 33.9% and 30.6% of the collocations were first introduced in textbooks at a level too difficult for the intended learners. From Japanese high school English textbooks, approximately 44.4% of the phrase collocations and 44.8% of the phrasal verb collocations were first introduced at a level that matched their collocation level. However, about 49.7% and 47.3% of the collocations were first introduced in textbooks at a level too difficult for the intended learners. In both the Japanese high school English textbooks and the CEFR-based textbooks, there was a general trend that as the collocation level increased, the proportion of Level Match collocations decreased, and that of Level Inappropriate collocations increased. This trend was observed for both phrase and phrasal verb collocations, with no significant difference between them.

From the comparison between Japanese high school English textbooks and CEFR-based textbooks, I found that the CEFR-based textbooks are more considerate of the timing of collocation introduction within reading comprehension sections. Additionally, a notable difference was found especially in phrase collocations compared to phrasal verb collocations.

Overall, although there are numerical differences, both high school English textbooks and CEFR-based textbooks may fail to fully consider the text levels when introducing collocations within reading comprehension texts. About only half of the collocations are introduced at a level matching the learners' proficiency, despite the critical role of the acquisition of collocations in second language learning. This suggests an urgent need for textbook improvement. Similarly, approximately one in three collocations in CEFR-based textbooks and one in two collocations in high school English textbooks were introduced to learners who are not proficient enough. These collocations are unlikely to be acquired by the learners due to their difficulty. Given the inherent constraints of English textbooks, such as limited themes and the number of words, it is challenging to repeatedly present the same

collocations. Consequently, the acquisition of collocations may rely heavily on their initial introduction. In this context, introducing collocations significantly above the learners' level within reading comprehension sections is unlikely to promote efficient learning.

There are some limitations to this study. First, it does not consider the frequency of collocations. This analysis treats collocations that appear only once and those that appear multiple times equally, assigning the text levels of their first appearances. It can be assumed that collocations targeted for acquisition are intended by the authors to appear multiple times, but this has not been considered, and all collocations have been weighted equally in the analysis.

Second, the study includes only the reading comprehension sections of the textbooks in the corpus. In textbooks, collocations are often taught not only through reading comprehension sections but also through grammar exercises and fill-in-the-blank questions to promote acquisition. However, these have not been included in the analysis.

Third, there is a bias in the collocations extracted for the study. This study analyzed collocations defined as two-word phrasal verbs and phrases in the EVP. However, phrases in EVP are limited to those whose meanings are difficult to infer from the meanings of the individual words. This reduced the number of collocations under consideration, particularly for A1 level phrasal verbs, of which only four were included. Consequently, the absence of a single collocation could affect the ratio by 25%, giving undue weight to individual collocations in the analysis. Although the EVP was used because it was the only available list of collocations with CEFR-assigned levels, future studies will require the development of a larger-scale level-specific collocation list that includes not only collocations whose meanings are difficult to infer from the individual words but also those that are easier to infer.

CONFLICT OF INTEREST

The author declares no conflict of interest.

REFERENCES

[1] S. Krashen and R. Scarcella, "On routines and patterns in language acquisition and performance," *Language Learning,* vol. 28, no. 2, pp. 283–300, 1978.

[2] A. Wray, *Formulaic Language and the Lexicon,* Cambridge University Press, 2002.

[3] J. Yamashita and N. Jiang, "L1 influence on the acquisition of L2 collocations: Japanese ESL users and EFL learners acquiring English collocations," *TESOL Quarterly,* vol. 44, no. 4, pp. 647–668, 2010.

[4] P. Lennon, "Getting 'easy' verbs wrong at the advanced level," *International Review of Applied Linguistics in Language Teaching,* vol. 34, no. 1, pp. 23–36, 1996.

[5] N. Nesselhauf, "The use of collocations by advanced learners of English and some implications for teaching," *Applied Linguistics,* vol. 24, no. 2, pp. 223–242, 2003.

[6] N. Nesselhauf, *Collocations in a Learner Corpus*, John Benjamins, 2005.

[7] S. Webb, J. Newton, and A. Chang, "Incidental learning of collocation," *Language Learning,* vol. 63, no. 1, pp. 91–120, 2013.

[8] A. Pellicer-Sánchez, "Learning L2 collocations incidentally from reading," *Language Teaching Research,* vol. 21, no. 3, pp. 381–402, 2017.

[9] J. C. Richards 2001. The role of textbooks in a language program. Retrieved February 5, 2023, [Online]. Available: https://www.professorjackrichards.com/wp-content/uploads/role-of-textbooks.pdf

[10] T. Takesue, "Collocation research based on junior-high school English textbooks in Japan," *The Journal of the Graduate School of Language and Literature,* vol. 12, pp. 125–137, 2019.

[11] T. Takesue, "Collocational analysis of the revised version of junior high English textbooks and a comparison with the previous version," *Journal of the Kyushu Academic Society of English Language Education,* vol. 50, pp. 41–50.

[12] R. Azemoto, "The analysis of collocations in high school English textbooks in Japan," *Journal of Corpus-based Lexicology Studies,* vol. 5, pp. 24–38, 2023.

[13] A. Doff, *Cambridge English Empower*. Cambridge University Press, 2015.

[14] L. Soars, J. Soars, and P. Hancock, *Headway,* 5th ed. Oxford University Press, 2019.

[15] P. Dummett, J. Hughes, and H. Stephenson, *Life*, 2nd ed. Cengage, 2018.

[16] D. Cotton, D. Falvey, and S. Kent, *New Language Leader*, 2nd ed. Pearson, 2014.

[17] S. Uchida and M. Negishi, "Assigning CEFR-J levels to English texts based on textual features," in *Proc. Asia Pacific Corpus Linguistics Conference,* vol. 4, 2018, pp. 463–467.

[18] R. Azemoto and S. Uchida, "The importance of criterion-referenced characteristics in CEFR-based English textbooks: Using CVLA indicators," *Studies in Language and Culture Studies,* vol. 48, pp. 35–47, 2022.

[19] Cambridge University Press. (2024). English Vocabulary Profile. [Online]. Available: https://www.englishprofile.org/wordlists/evp

[20] Ministry of Education, Culture, Sports, Science and Technology (MEXT). (2020). The Relationship between the High School Curriculum Guidelines and English Examinations/Proficiency Tests. [Online]. Available: https://www.mext.go.jp/content/20200318-mxt_daigakuc02-000005103_5.pdf