論文

TOEIC Vocabulary and EIKEN Vocabulary Compared Stephen M. Ryan¹⁾

Keywords: TOEIC, EIKEN, Vocabulary

Students of English in Japan have a wide choice of standardised tests they can use to demonstrate their English proficiency to themselves, their schools, and potential employers. By far the most popular among these are the Test of English for International Communication (TOEIC), made by the Educational Testing Service (ETS) and the Test in Practical English (実用英語技能検定), usually known as EIKEN, run by the Eiken Foundation of Japani.

Both tests aim to give a comprehensive assessment of a student's English proficiency, and, superficially, appear to be targeting a similar kind of English. The distinction between "English for International Communication" (TOEIC) and "Practical English" (EIKEN) would appear to be small. Even when TOEIC claims to be "a global standard for assessing English communication skills objectively and fairly" (Institute for International Business Communication, n.d.) and EIKEN describes itself as "Japan's national English test" (Eigo Kentei Kyoukai, n.d.-a), the difference does not seem too large. Indeed, universities and language skills often run classes to prepare students for both test at once, under the umbrella title of "Qualification [Test] English" (e.g. Sanyo Gakuen Daigaku, 2019).

However, a cursory glance at past papers from the two tests shows that their focus is very different. TOEIC is clearly designed to assess students' ability to use English in the business world. Many of its questions involve business contexts; reading texts are often invoices, orders, advertisements, or other business communications. A pamphlet issued by ETS to promote the test (2009) calls it "the global standard for assessing English proficiency in the workplace." EIKEN makes no similar claim. It says its target is "practical English proficiency" (Eigo Kentei Kyoukai, n.d.-b), but many of the questions, especially in the Reading section, appear to have an academic focus, resembling journalistic reports on scientific research.

Clearly, the English required to pass a test on Business English can be quite different from that required to understand journalistic/scientific reports. There will be differences of genre, discourse features, register, vocabulary, and relevant background knowledge. Students preparing for these tests, however, tend to focus on vocabulary: if only they knew more English words, they seem to believe, their scores on "Qualification Tests" would improve.

¹⁾ 山陽学園大学総合人間学部言語文化学科

This investigation, then, will focus on the vocabulary used in TOEIC and EIKEN tests, asking to what extent there is a core vocabulary shared by both tests and to what extent Business English for TOEIC and journalistic/scientific language is needed for EIKEN. The answer should be of practical significance to teachers and students involved in preparing for both tests at the same time or in the same lesson.

Previous Research

There is no shortage of studies comparing TOEIC and EIKEN, sometimes with and sometimes without other English proficiency tests. However, these comparisons have focussed mainly on the equivalence or otherwise of scores obtained on the various tests (CASEC, n.d.; Eigo 4 Ginou Shikaku Kentei Shiken Kondan-kai, 2016; *Eiken TOEIC Shiken*, n.d.; Hirai, 2017), or on the suitability of one or the other test for placing students in specific programmes (Innami & Koizumi, 2017).

Studies looking specifically at the vocabulary used in these tests have focussed on one test or the other (Chujo & Genung, 2005 - for TOEIC; Miura & Beglar, 2002 - for EIKEN), without making overt comparisons between the two. One notable exception is Chujo and Oghigian (2009), which looked at TOEFL as well as TOEIC and EIKEN and asked how many words a test-taker needed to know in order to understand each, by counting types and tokens in the Reading sections of two EIKEN pre-1st level testsii and two TOEIC tests, and comparing their data with the High Frequency Wordlist of the British National Corpus (http://www.natcorp.ox.ac.uk/using/index.xml?ID=freq). Their conclusion, that a significantly higher level of vocabulary knowledge is needed for EIKEN Pre-1st than for TOEIC is hardly surprising, given that EIKEN Pre-1st is designed for students with higher levels of English proficiency than the average of those taking TOEIC. Eigo wo Raku ni Manabitai (2020) also presents data on the overall size of vocabulary needed to pass each of the different levels of the EIKEN.

This investigation will focus on a direct comparison of vocabulary needed for TOEIC and EIKEN.

Methodology

The Listening and Reading sections of three tests were analysed for this investigation. The tests were chosen at random from collections of past papers (IIBC, 2016; Sebido, 2019a, 2019b; TOEIC, 2016) which were available to the researcher and all come from the years 2016ⁱⁱⁱ-2018. There was one TOEIC test (Test 1, 2016), one EIKEN 2nd level test (Test 2, 2018), and one EIKEN Pre-1st level test (Test 1, 2018).

Analysis of two different levels of EIKEN was necessitated by an important difference between TOEIC and EIKEN: whereas TOEIC offers a single test for all students, awarding scores that indicate proficiency levels, EIKEN offers seven different levels of test (from low to high: 5, 4, 3, pre-2, 2, pre-1, 1) on a pass/fail basis. A TOEIC test then covers a number of different EIKEN levels. Levels 2 and pre-1 of EIKEN were chosen for analysis as the levels

of these two tests correspond to TOEIC scores obtained by the majority of students taking the test: roughly 550 to 940 on a scale of 10-990 on the combined Reading and Listening TOEIC (*Eiken TOEIC Shiken*, n.d.).

For Reading sections, all English words printed in texts, questions, and potential answers in the testing booklet were included in the analysis. For the Listening section, not only the printed words in the testing booklets but also the English words in transcripts of the recording were used.

A word falling into one of the following four categories as noted and its frequency in the test was analysed: Nouns, Verbs, Adjectives, Adverbs. A fifth category, Adjectivals (consisting of multi-word combinations with adjectival force) was abandoned after it proved rather impervious to lexical analysis. Proper nouns were excluded, except in the case where the noun was also a common noun: so, "Richmond Park" was counted as one for "park." The words and their frequencies were then tabulated and the tables for the three tests placed side by side for analysis.

Results

A total of 3,355 words were observed, including 1,802 different nouns, 739 verbs, 618 adjectives and 199 adverbs. The totals of unique words in each test for each category are shown in Table 1. A full listing of all these words cannot be included here because of space limitations. However, the list, as well as lists showing words in each of the categories for which numbers are given in Tables 1 and 2, is available on request from the authoriv.

Table 1Total Number of Different Nouns, Verbs, Adjectives, and Adverbs in Each Test

	Nouns	Verbs	Adjectives	Adverbs	TOTAL
EIKEN 2	445	254	179	87	965
EIKEN pre-1	719	419	330	114	1,582
TOEIC	1,084	436	330	121	1,971

Table 2 shows a further breakdown of the total number of unique words into those that occur on all three tests; those on both the Level 2 EIKEN and TOEIC; the Level Pre-1 EIKEN and TOEIC; and, finally, those occurring only in TOEIC and only in EIKEN. Numbers in parentheses express the word count as a percentage of all unique words in that grammatical category for all three tests. Thus, the 83 nouns that occur in all three tests represent 4.6% of the 1,802 different nouns that occur in the three tests.

Table 2

Totals and Percentages of Unique Words Occurring in All Three, Two, and Only One of the Tests

	Nouns	Verbs	Adjectives	Adverbs	TOTAL
Total of unique words	1,802	739	615	199	3,355
Words that occur in all three tests	83	94	48	38	263
	(4.6)	(12.7)	(7.8)	(19.1)	(7.8)
Words that occur in both EIKEN	143	135	74	50	402
Level 2 and TOEIC ^a	(7.9)	(18.3)	(12.0)	(25.0)	(12.0)
Words that occur in both EIKEN	202	170	99	49	520
Level Pre-1 and TOEIC ^a	(11.2)	(23.0)	(16.1)	(24.6)	(15.5)
Words that occur only in TOEIC	784	209	192	48	1,233
	(43.5)	(28.3)	(31.2)	(24.1)	(36.8)
Words that occur only in the EIKEN tests ^b	725	302	288	85	1,400
	(40.2)	(40.9)	(46.8)	(42.7)	(41.7)

Note. Percentages (in parentheses) were obtained by dividing the word count in each cell by the total number of unique of words in each grammatical category (given in the first row of this table).

Analysis

The fact that only 7.8% of all words (nouns, verbs, adjectives, and adverbs) occur on all three tests is not good news for test takers seeking to learn words for both EIKEN and TOEIC at the same time. There is more overlap between each of the levels of EIKEN analysed and the TOEIC: 12.0% shared at Level 2 and 15.5% at Level Pre-1. However, none of these percentages is particularly large. This gloomy situation is further illuminated by the fact that 36.8% of the words are unique to TOEIC and 41.7% to EIKEN (when totals for both levels of EIKEN are combined.

The overlap between the three tests is greatest in the case of adverbs: 19.1% of all adverbs occur on all three tests. Since the total number of adverbs in all tests is only 199, this suggests that time spent studying adverbs will result in a better understanding of both TOEIC and EIKEN tests. This approach, however, applies least in the case of nouns, of which only 4.6% occur on all three tests.

The comparison of the TOEIC vocabulary with each of the two different levels of EIKEN gives a more detailed understanding of the shared words. The 12.0% of all words that appear in both TOEIC and Level 2 EIKEN can be considered the basic vocabulary for taking either of these tests. Again, it is adverbs that show the most overlap and nouns the least.

^aIncludes words occurring in all three tests.

bBoth EIKEN (Level 2 and Level Pre-1) are included in this statistic.

Beyond the basic level, there is more overlap, 15.5% of words on the EIKEN Level Pre1 test also appearing in the TOEIC test. There may be some amount of comfort here for
learners: the more their English progresses the more overlap there is in the vocabulary
needed for the two tests. However, 15.5% is not a large amount, especially when Hatori
(1979) considers that, in order for a reader to understand a written text, 19 out of every 20
words should be known to the reader (the so-called "95% coverage").

Although adverbs continue to show the most overlap (24.6%) and nouns the least (11.2%) in the TOEIC / EIKEN Pre-1 comparison, verbs also show a significant level of overlap (23.0%), suggesting that as students' level of English proficiency increases from EIKEN 2 to EIKEN Pre-1, they should be advised to shift their focus to include more verbs in their vocabulary study.

The focus on specific word categories (nouns, verbs, adjectives, and adverbs) in the analysis to this point is prompted by the fact that students (in Japan, at least) tend to view, and study, words independently from each other, considering each word to be a separate item that needs to be understood and memorised in isolation from each other. The prevalence of test preparation books which are essentially lists of words, with their translation and examples of their usage (eg. Gakken, 2018; Kato, 2019; Oubunsha, 2012) tends to reinforce this habit.

However, a basic understanding of the relations between words can be a powerful multiplier for vocabulary learning. If students understand that words occur in families of related nouns, verbs, adjectives, and adverbs, with often visible/audible similarities and regularities (eg. a base, to be based, basic, basically), their learning burden is considerably reduced. It is probably more important for the teacher to inculcate an understanding of these "family" relations among words than to stipulate exactly which words need to be learned.

This leaves the troublesome issue of the 36.8% of words that are unique to TOEIC and the 41.7% that are used only in EIKEN. Some of these are words that even very advanced users of English could not be expected to know and they fall clearly into the 5% of words in a text that Hatori (1979) says are not needed for a student to understand a text. Such words include "festivalgoer" from TOEIC, "cytokine," and "cadmium" from EIKEN. These very low frequency words were more likely to be found in the EIKEN than the TOEIC dataset, perhaps because of the tendency of EIKEN to include texts dealing with scientific and technical topics. Frequency within the dataset was not found to be a good guide to identifying these words: "cadmium," for example, occurs four times in the EIKEN dataset. This is because once a word occurs in a text it can be repeated in that text or in the questions and suggested answers that follow it.

Students can be best equipped for dealing with such words by developing both an ability to guess the meaning of a word based on the context in which it occurs and a number of "word attack" skills. Being able to guess more or less what a word means, or at least what role it performs in a sentence, by taking cues from the surrounding text and knowledge of the world is a skill which will help students not only with extremely low-frequency words but also with higher frequency ones which they have not encountered before. Good

readers/listeners will already have this skill in their native language, but may need encouragement to transfer the skill to English. Word attack includes identifying the part of speech of a word, breaking it down into potential prefixes, suffixes, and stems, and trying to recall cognates or similar looking/sounding words encountered previously.

A further source of divergence between TOEIC and EIKEN datasets was the researcher's decision to view noun clusters (such as "assembly line" from TOEIC or "franchise owner" from EIKEN) as a single noun and count it independently of its component nouns. Such clusters account for a significant proportion of the nouns in the datasets (43.1% of nouns unique to TOEIC and 24.3% of unique EIKEN nouns) and can be both long and complex: "metro area business watch" (TOEIC), "navigation software market" (EIKEN). Although the combinations may be complex, more often than not the clusters are made up of high-frequency words that should be familiar to students taking the test: "membership offer," "mall management" (TOEIC), "computer science professor," and "car body shop" (TOEIC). What is challenging here is not the words themselves but the way in which their meanings can combine, knowing the difference, for example, between a "volunteer parent" (EIKEN) and a "parent volunteer," or between "conference materials" (TOEIC) and a "materials conference." The ways in which nouns interact in a cluster can be taught and will, again, serve as vocabulary multiplier. Learning them is a far more achievable goal than learning all the possible combinations of familiar nouns.

The prevalence of noun clusters is perhaps not surprising in texts where focusing, respectively, on business (TOEIC) and science/technology (EIKEN), both fields notorious for grouping together nouns in long strings. What is surprising is the proportionately greater number of numbers to be found in TOEIC compared to EIKEN.

What, then, of the contention that students preparing for TOEIC need to focus on business vocabulary and those taking EIKEN on science/technology terms? Analysis of the 50 most frequent words (nouns, verbs, adjectives, and adverbs) on each test reveals eight overtly business-related words on TOEIC (front desk, supervisor, subscriber, production, meeting room, invoice, internship, business owner), but only three (disease, professor, test) science/technical words on the EIKEN. The next most frequent show a similar picture: six overtly business terms in TOEIC and four vaguely science/technical words in EIKEN, although EIKEN also has the business term "franchise."

There are certainly more business-related terms in TOEIC than EIKEN: the word "business" itself occurs only twice in the EIKEN dataset, but twelve times in the TOEIC data. So, it would seem that there is merit in having students preparing for TOEIC learn business-related terms. However, their prevalence is not as strong as one might expect from the large number of documents from a business setting appearing in the TOEIC test. It may be just as important to focus on an understanding of business genres and business discourse as it is to spend time on the words themselves.

On the other hand, an EIKEN bias in favour of science/technology is barely discernible: the stem "scien-" occurs only twice in the EIKEN data but once in the TOEIC data. There does not seem to be much evidence that time spent studying

scientific/technological words would be well spent by those preparing for EIKEN. As with TOEIC, though, familiarity with the appropriate genre and discourse patterns (in this case, those of journalistic reports on scientific matters) would be more useful than knowing the words themselves.

In both tests, however, the most frequently used words are usually words from daily life, rather than those with particular business or science/technology associations, although there are some bizarre exceptions. Table 3 shows the ten most frequent words on both tests. The extraordinary prevalence of both "mummy" (preserved human remains, rather than maternal parent) and "banana" arise from two reading texts, one about mummies and the other about bananas. This suggests strongly that the dataset used for this study was not large enough to mitigate the effects individual texts. In other words, it does not capture the true frequency of vocabulary items in TOEIC or EIKEN (2, pre-1) tests as a whole. A much larger dataset will be needed to do this.

Table 3
Top Ten Most Frequent Words on the Tests

TOEIC		EIKEN ^a		
Word	Frequency	Word	Frequency	
	count		count	
workshop	14	mummy	21	
medicine	14	student	17	
indicate	14	other	17	
date	13	eye	13	
phone	12	banana	13	
mall	12	wear	12	
notice	11	study	12	
item	11	son	12	
schedule	9	safe	12	
mention	8	effect	12	

Note. ^aThese counts include both Level 2 and Level Pre-1.

The presence of "indicate" as the third most frequent word in the TOEIC test studied is noteworthy. Its prevalence comes from the use of a particular question type, which asks for example, "What does the woman indicate about her plans?" This question type is often used in the Listening part of the test, in preference to questions that use words that occur more frequently in daily life, such as "say" or "suggest." A student who did not know the word "indicate" would be at serious risk of under-performing despite having understood enough of the foregoing conversation to answer the question. Thus, there would seem to be value in studying not only the types of questions that are asked on the tests but also the

specific words that are used to ask them. Even if those words occur only infrequently in general English usage.

The real problem, though, for students taking the test is not the words that occur most frequently on the tests, like "medicine," "date," "student," and "other," which they tend to know from the early days of their study of English. The real problem is the words that occur infrequently in the tests but can be crucial to understanding a text or answering a question. The number of words that occurs only once in the datasets is huge: 67.8% of the words unique to EIKEN and 71.6% of the words unique to the TOEIC. What is the learner to do about these words (1,832 in total, across the two tests, or 1,740 if we take on board Hatori's, 1979, 95% coverage for full comprehension)? It is not realistic to imagine a student sitting down to memorise all of these words, plus the many more words that *might* have come up, in preparation for the test. Word attack skills and guessing from context will get the student only so far. What is to be done about these words?

The answer to this question is as straightforward as it is transparent. Both sets of test makers mention it in their tag-line statements of purpose: TOEIC is "the global standard for assessing English proficiency in the workplace" (ETS, 2009); EIKEN targets "practical English proficiency" (Eigo Kentei Kyoukai, n.d.-b). The tests are not designed to measure who has learned the most English words; they are designed to measure English proficiency. The best way for students to ensure they know the words necessary to understand and pass the test is to improve their English proficiency. They should study the language, expose themselves to a lot of it, ask questions, take every opportunity to use it, invite feedback, etc., etc. Memorising vocabulary items from a carefully prepared list (eg. Gakken, 2018; Kato, 2019; Oubunsha, 2012) is an attractive approach; many of the focusses suggested in this paper can be helpful; but, in the end, they are no substitute for studying the language as a whole and improving proficiency.

Conclusion

This study began by asking if there were enough similarities between the vocabulary of TOEIC and the English of EIKEN to justify having students prepare for both tests in the same way. It has revealed that the quantity of vocabulary shared by the two tests is surprisingly small and that TOEIC does tend to use business vocabulary more frequently than EIKEN. It has led to a number of suggestions that might help mitigate these differences:

- A focus on adverbs in test preparation for less proficient students
- A gradual shift to focussing on verbs as students become more proficient
- Cultivation of the ability to deduce meaning from context
- Use of a number of word-attack skills
- Study of the ways in which nouns combine to make clusters with specific meanings
- Exposure to the specific genres and discourse types used most often in these tests:
 business for TOEIC and journalistic reports on science/technology for EIKEN

However, implementation of all of these suggestions in an ideal test preparation class would still leave a huge number of words that students need to know in order to pass either of these tests. The only certain way to ensure success on TOEIC or EIKEN is to help students to increase their English proficiency.

In the end, then, the answer to the question posed at the start of this study is "Yes." Students can prepare for both tests at the same time and in the same class by working on their English proficiency, the very thing the testers set out to measure. There are some short-cuts and techniques they can use (listed above) but the heart of test preparation needs to be increasing English proficiency.

References

- CASEC. (n.d.). Eigo komyunikeeshon nouryoku hantei tesuto [English communication ability test]. https://casec.evidus.com/materials/
- Chujo, K., & Genung, M. (2005). *Utilizing the British National Corpus to analyze TOEIC tests: The quantification of vocabulary-usage levels and the extraction of characteristically used words.* Tokyo, Japan: Institute for International Business Communication. http://www.iibc-global.org/library/default/english/lr/about/data/research/pdf/3_E.pdf
- Chujo, K., & Oghigian, K. (2009). How many words do you need to know to understand TOEIC, TOEFL and EIKEN? An examination of text coverage and high frequency vocabulary. *The Journal of Asia TEFL*, 6(2), 121-148.
 - https://pdfs.semanticscholar.org/f7dc/b4fa90d82f489b22415db1dafdfe43cd457d.pdf
- Educational Testing Service. (2009). Fair. Valid. Reliable.: Qualities that make TOEIC tests the global standard for assessing English proficiency in the workplace. Princeton, NJ: ETS.
- https://www.amideast.org/sites/default/files/otherfiles/hq/advising%20testing/toeic_quality_brochure-low_res-10194_0.pdf
- Eigo Kentei Kyoukai [Society for Testing English Proficiency]. (n.d.-a). *EIKEN tests*. https://www.eiken.or.jp/eiken/en/eiken-tests/
- Eigo 4 Ginou Shikaku Kentei Shiken Kondan-kai [English 4 Skills Qualification and Examination Roundtable]. (2016). *Shikaku ekntei shiken CEFR to no taishō-hyō* [Comparison table of qualification examinations and the CEFR]. Eigo 4 Ginou Shiken Joho Saito [English 4 Skills Examination Information Site]. http://4skills.jp/
- Eigo Kentei Kyoukai [Society for Testing English Proficiency]. (n.d.-b). *Overview of the EIKEN tests.* https://www.eiken.or.jp/eiken/en/eiken-tests/overview/
- Eigo wo Raku ni Manabitai [I Want to Study English with Ease]. (2020). Eiken 1 ~5-kyū no goi-sū, eitango reberu no hikaku [Comparison of the number of English words for levels 1-5 of Eiken. https://ei-raku.com/2017/11/exam-eiken-comparison/
- Eiken TOEIC Shiken [EIKEN and TOEIC Tests], n.d.). http://mellow.na.coocan.jp/tests.htm

- Gakken (Ed.). (2018). *Eiken 2-kyuu eitango 1750 eiken ranku-jun* [EIKEN level 2 1750 English words, ranked in EIKEN order]. Tokyo: Gakken.
- Hatori, H. (1979). *Eigo shidouhou handbook (4) Hyouka-hen* [A handbook for English teaching (4) evaluation]. Tokyo: Taishukanshoten.
- Hirai, M. (2017). A general overview of English tests administered in Japan. *Shiken* [Examinations], *21*(1), 12-22.
- http://teval.jalt.org/sites/teval.jalt.org/files/21_01_12_Hirai_Overview_of_Tests.pdf IIBC (Ed.). (2016). *Koushiki TOEIC listening & reading: Mondaishu 1* [Official TOEIC listening and reading: Collection of papers 1]. Tokyo: IIBC.
- Innami, Y., & Koizumi, R. (2017). Using EIKEN, TOEFL, and TOEIC to award EFL course credits in Japanese universities. *Language Assessment Quarterly*, 14(3), 274-293. https://www.tandfonline.com/doi/abs/10.1080/15434303.2016.1262375
- Institute for International Business Communication. (n.d.). *About the TOEIC program*. https://www.iibc-global.org/english/toeic/toeic_program.html
- Kato, T. (2019). *TOEIC L&R TEST deru-tan tokyuu: Kin no sentensu* [Super express to the words on the TOEIC listening and reading tests: Gold sentences]. Tokyo: Asahi Shinbun.
- Miura, T., & Beglar, D. (2002). The Eiken vocabulary section: An analysis and recommendations for change. *JALT Journal*, *24*(2), 107-129. http://jalt-publications.org/files/pdf/jalt_journal/2002b_jj.pdf#page=7
- Oubunsha (Ed.). (2012). *Deru-jun pasu-tan: Eiken jun-1-kyuu* [Words you need to pass, appearing in order: EIKEN pre-1 level]. Tokyo: Oubunsha.
- Sanyo Gakuen Daigaku [Sanyo Gakuen University], (2019). *Kougi Gaiyou: Shirabasu* [Course Outlines: Sylabuses]. Okayama, Japan.
- Seibido Publishing Editorial Department (Ed.). (2019a). *Eiken jun-1-kyu: kako6kai mondaishu, 19nendo* [EIKEN level pre-1: Collection of papers from the last six times, 2019]. Tokyo: Sebido.
- Seibido Publishing Editorial Department (Ed.). (2019b). *Eiken 2-kyu: kako6kai mondaishu, 19nendo* [EIKEN level 2: Collection of papers from the last six times, 2019]. Tokyo: Sebido.

ⁱ Formerly known as the Society for Testing English Proficiency.

ii See below for an explanation of EIKEN levels.

iii Significant changes were made to TOEIC in 2016. The test analysed incorporated those changes.

iv ryan@sguc.ac.jp