

Cross-Sectional Analyses of TOEIC Vocabulary

Natsue NAKAYAMA

1. Introduction

The Test of English for International Communication or TOEIC has recently drawn wide attention in communication-oriented English education in Japanese universities as a form of assessment that can judge the communicative competence of learners. In fact, the Institute for International Business Communication (or IIBC) states, there were 442 universities which adopted TOEIC in 2005. Considering the fact that there were 726 universities in Japan when the above mentioned data was collected (Ministry of Education, Culture, Sports, Science and Technology), we can see the high percentage of adoption of TOEIC at universities. One reason for the wide recognition of this test is that companies attribute a certain amount of significance to it. "Some companies are even using the TOEIC score as one of the criteria for an employee's promotion" (IIBC 20). Seeing the wide popularity of this test among the companies, universities began to train their students in strategies for taking the TOEIC. Moreover, in order to prepare their students to be more competitive in future job hunting, some universities even considered their students' TOEIC score as a part of assessment towards acquiring university credits.

One of the characteristic features of the TOEIC pertains to the sizeable range of vocabulary that is used in it. In addition, according to Ishikawa (201-214), the TOEIC examinees are aware of the fact that this vocabulary is rather specialized. In fact, as Nation & Newton (238-254) argue, every field has its own technical vocabulary, which is used within a narrow range and is normally limited to a specialized field. Keeping these points in mind, in order to prepare learners for this test, it appears worthwhile to examine what kind of vocabulary is used in the TOEIC and determine the characteristic features of 'TOEIC vocabulary'.

Many of the materials that are published for learning TOEIC vocabulary list words according to the frequency of their appearance in the test as a whole. In comparison, there are fewer books that list words according to the sections that they appear in (whether the words appear more frequently in the listening or reading section). Nation (27) explains that there are 18 aspects involved in knowing a word. According to his classification of word knowledge, the aspect that is tested in the listening section of the TOEIC can be classified as 'receptive knowledge' regarding 'what the word sounds

like', which is listed under 'form'. On the other hand, the word knowledge that is tested in the reading section can be classified as 'receptive knowledge' regarding 'what the word looks like', which is also listed as a subcategory of 'form'. From this, we can assume that the listening and reading sections of the TOEIC test different aspects of words. If different aspects of words are tested based on the respective sections—either listening or reading—we can assume that the vocabulary used in one section will differ from that in the other. However, few studies examine TOEIC vocabulary and its features on the basis of the sections in which it appears.

This study attempts a cross-sectional investigation of TOEIC vocabulary. As Tanimura & Utiyama (81-96) summarize, vocabulary for specific areas should be selected using appropriate corpus data. Accordingly, a small range of vocabulary was collected so as to build a mini-TOEIC corpus of the listening and reading sections. By conducting a comparison of the findings, I would like to identify the vocabularies that are common to each section, listening or reading, and their respective features. The findings are expected to be of use in preparing a more focused approach for the learners aspiring to take the TOEIC.

2. Background

2.1 Defining 'basewrd'

This study utilizes the software 'Range 29b¹' developed by the University of Victoria to analyse the word level and frequency of TOEIC vocabulary. Range 29b can apply three distinct word lists, i.e. base word lists, to text data and sort the text vocabulary into three categories from each list. It also lists words from outside the three lists, thus resulting in the formation of a total number of four categories. In the present study, the following three base word lists will be used as criteria for the word level analysis of TOEIC vocabulary.

basewrd 1 = General Service List (GSL)'s first 1000 words by West

basewrd 2 = GSL's second 1000 words (or 2000 words)

basewrd 3 = Academic Word List (AWL) words by Coxhead

The words that are used in these lists are referred to as 'basewrd' instead of 'baseword'. This is to reflect the term that is used in the Range 29b software program.

2.2 How to count words

Range 29b is a software that counts the words that are used in a text data in three different ways:

(1) the number of running words used in a text (or tokens), (2) the frequency of a headword (or types), and (3) a family frequency figure (or word families). Of the three

ways of counting, 'word family' counts a headword and its inflected and derived forms as one word. Due to this feature, counting words on the basis of their word family is likely to significantly reduce the number of words considered, as compared to counting words as tokens or types. The objective of this study is to examine the features of the words that are used in the listening and reading sections of the TOEIC; for this, it appears to be easier to gain a general overview of the words used when their numbers are smaller. This is accomplished by considering derivative words as one word. Thus, this study adopts 'word family' as a method of counting words.

Furthermore, due to the limitation of the software used in this study, homonyms were classified as the same word, despite their meanings or the parts of speech that they constituted. In addition, words that did not appear in the basewrd lists or not in the list words (hereafter, NITL) were counted in types due to the following two reasons. First, the lemmatization of these words was not possible through this software. Also, the words that appeared in the list were low frequent words so that the Japanese learners of English whose L1 is unrelated to English, had a possibility of not being able to recognize them even though they knew the base words.

3. Study

3.1 Purpose

The purpose of this study is to analyse TOEIC vocabulary from the viewpoint of frequency and word level and to discover the characteristic features of the vocabulary that is used in the TOEIC's listening and reading sections. The following four sub-purposes were formulated in order to attain this purpose:

- (1) Determine the distribution of the frequency and word level of TOEIC vocabulary according to the section in which it appears.
- (2) Determine the rate of coverage of the shared vocabulary that is used in both the sections.
- (3) Extract the 'high frequency words' that appear in each section and discover the distribution of their respective word levels.
- (4) From the high frequency vocabulary, identify the word items that are commonly used in both the sections and their corresponding percentages.

3.2 Materials

The mini-TOEIC corpus was compiled from the data extracted from five sets of sample TOEIC papers and two sets of exercises on preparation for the listening and reading sections from the Official TOEIC Guide & Exercise Book, vols.1 & 2 by IIBC. Compiling a corpus from two source books is likely to limit the data that are studied.

However, in light of the fact that these are the only two official books published by the TOEIC test makers², it is necessary to limit the data in this study so as to maintain accuracy in the corpus data.

All the exercises—including the questions and the alternatives associated with the answers—were scanned while creating the corpus. However, the instructions and question numbers in each part were excluded from the data because they were considered to act as signs.

3.3 Procedure

An electrical corpus of TOEIC vocabulary was initially created. Range 29b was executed on this corpus to analyse the sectional frequency and word level distribution of TOEIC vocabulary. The words that were not included in any of the three base word lists (or NITL words) listed above were further classified into three categories: a, b, and c. Based on the results of this classification, the number of NITL words was modified according to the following formula: $NITL (type) = \text{pre-modified NITL (type)} - a - b - c$.

The process of modifying the number of NITL words was followed in accordance with Aizawa & Ono. During this process, the words were classified into either a, b, or c on the basis of the following criteria: (a) a proper noun, (b) a commonly known word, and (c) an abbreviation. Of the three, it was possible for criteria b to be subjectively affected by the person who judged the classification of the words. Thus, further criteria were set for criteria b: (1) English loan words that Japanese learners of English were familiar with as Japanese words and (2) words that did not occur in the GSL (base word lists) but were frequently used in Japanese EFL context—in other words, the words that were listed as the 1000 most frequent words in JACET 8000.

Moreover, in order to clarify which words were used in both the listening and the reading sections, words used in the reading section were set as *basewrd*, and the rate of vocabulary coverage in the listening section was calculated.

Following this, the high frequency words in each section were extracted. Due to the limitations of this paper, the top 60 most frequently occurring words are defined as ‘high frequency words’ in this paper.

Finally, the high frequency words that appeared in the reading section were set as *basewrd*, and the coverage of vocabulary with the high frequency words of the listening section was calculated based on this list. From this, I hoped to identify the high frequency words that commonly appeared in both the sections.

4. Results

(1) Determining the sectional frequency, word level, and the distribution of the word

levels in TOEIC vocabulary

Table 1 summarizes the sectional frequency of the occurrence and word levels of TOEIC vocabulary. Table 2 shows the sectional distribution of NITL and the modified result of NITL. Table 1 indicates that 9.73% of the words in the listening section could be classified as AWL, while AWL words constituted 14.24% of the reading section vocabulary. Table 2 shows that in the listening section, 11.26% of the words could be classified as modified NITL, whereas in the reading section, this figure was at 17.95%. In short, 20.99% of the words in the listening section and 32.19% of those in the reading section could be listed as low frequency high level vocabulary words such as AWL and NITL words. This indicates that a higher level of vocabulary is used in the reading section than in the listening section.

One of the features of the TOEIC vocabulary in the listening section is the extensive use of high frequency words. The results indicate that high frequency lower level words such as 1000 word level vocabulary and category b NITL words are more frequently found in the listening section than in the reading section. Table 1 illustrates that the 1000 level words constitute 42.66% of the listening section as compared to 33.77% of the reading section. Table 2 indicates that category b NITL words constitute 10.53% of the listening section, but only 7.01% of the reading section. As regards proper nouns, they were found to be more prevalent in the reading section: they constituted 12.97% of the reading section, but only 7.89% of the listening section.

Table1. Sectional Frequency of the Occurrence and Word Levels of TOEIC Vocabulary

section	tokens		types		Word families		Function words	
Listening	number	[%]	number	[%]	number	[%]	number	[%]
1000 word	10636	83.35	1208	51.4	697	42.66	125	10.93
2000 word	846	6.63	380	16.17	288	17.63	0	0.00
AWL	391	3.06	221	9.40	159	9.73	1	0.09
NITL	887	6.95	541	23.02	490	29.99	0	0.00
sum	12760	100	2350	100	1634	100.00	126	11.01
section	tokens		Types		Word families		Function words	
Reading	number	[%]	number	[%]	number	[%]	number	[%]
1000 word	10657	73.26	1413	41.57	747	33.77	132	9.70
2000 word	1037	7.13	443	13.03	288	13.02	2	0.10
AWL	1209	8.31	541	15.92	315	14.24	3	0.20
NITL	1644	11.30	1002	29.48	862	38.97	0	0.00
sum	14547	100	3399	100	2212	100.00	137	10.10

Table 2. The Sectional Distribution of NITL and the Modified Result of NITL

Listening		Words (types)	% [running words/ types]	% [types// pre-modified NITL sum]	Reading		Words (types)	% [running words/ types]	% [types// pre-modified NITL sum]
N I T L	A	129	7.89	26.33	N I T L	A	287	12.97	33.29
	B	172	10.53	35.10		b	155	7.01	17.98
	C	5	0.31	1.02		c	23	1.04	2.67
	Modified NITL	184	11.26	37.55		Modified NITL	397	17.95	46.06
	NITL sum	490	29.99	100.00		NITL sum	862	38.97	100.00

(2) Text coverage of the TOEIC vocabulary in the listening and the reading sections

Table 3 shows the rate of coverage of the vocabulary used in the listening section by the reading section vocabulary. The results indicate that 860 words were commonly used in both the sections, which constitutes 52.63% of all the words that were used.

Table 3. Text Coverage of the TOEIC Vocabulary in Both the Sections

	Word family	Common words in both section / Whole running words of the listening section
Rate of coverage with the Reading section	860	52.63%
N I T L	774	47.37%
Sum	1634	100%

(excluding function words)

(3) Sectional high frequency words (top 60 words) and their features

Appendix 1 lists the top 60 words that occurred most frequently in the listening and reading sections. Table 4 summarizes the respective features of the top 60 words from each section. These features include the frequency rate of the first 1000 level words. As is evident from the table, of the top 60 most frequently used words, 1000 level words comprised 96.67% (58 words) of the listening section and 85.00% (51 words) of the reading section.

In this section, further analysis was conducted to determine the features of the sectional top 60 words. Since business is generally assumed to be one of the most popular subject areas for the TOEIC exams, 'business terms' were counted in order to gauge how many of the top 60 words were from this category. In order to extract business terms from the top 60 words, TOEIC vocabulary list 1, 2, & 3 by Chujo (27-42) and Chujo et al. (29-43) was used as a base word lists. These are online word lists used in the custom-made software, by Chujo et al, to prepare their students for the TOEIC. There are two lists in this software: TOEIC vocabulary 1 & 2 (Chujo, 2003)

which lists words that aims to build around 200-300 TOEIC score for the level 1, and 300-400 for level 2, and TOEIC vocabulary 3 (Chujo et al, 2004) which lists words that aim to build 400-500. The procedure of extracting business terms was as follows. First, of the words that are listed in the 'TOEIC vocabulary list 1, 2 & 3', those that are classified under the business related topics such as business, office, marketing, human resource, and purchases were defined as 'business terms' in the present study. Subsequently, the types of words that appear in this list were lemmatized and listed according to their word family. This procedure was necessary to unify the unit of the counted words in both the lists in order to derive the rate of coverage between the top 60 words and the business terms. Finally, the coverage rate was calculated. The findings regarding the business terms in the sectional top 60 word lists and their features are depicted in Appendix 1 and Table 4. As is evident in Table 4, business terms appear more frequently in the reading section than they do in the listening section. More specifically, while business terms account for 8.33% or five words of the top 60 words in the listening section, they constitute 25.00% or 15 words of the top 60 words in the reading section.

Finally, the top 60 words were indeed found to be high frequency words of the TOEIC vocabulary, and it was demonstrated that the top 60 words from the listening and reading sections comprised over 11.81% of all the words in the mini-TOEIC corpus created for this study.

Table 4. Features of the Sectional Top 60 Words

Section	1000 word	[1000 word/60words]	Business terms	[Business terms/60words]	Frequency of appearance	[Frequency/sectional running words]	[Frequency/Whole running words]
Listening	58	96.67	5	8.33	1526	11.96	11.81
Reading	51	85.00	15	25.00	1699	11.68	

(4) Text coverage of the top 60 words of the listening and the reading sections

Table 5 shows the rate of coverage of the top 60 words in the reading section by those in the listening section. The results of the calculation indicate a tendency similar to the findings in 4.1. In other words, a similar tendency is discerned irrespective of whether the comparison is made between the sections of the entire list of running words or only the top 60 words. It was clear that in both cases, nearly half of the vocabulary was common to both the sections. As regards the top 60 words, 26 words (43.33%) were common to both the sections. The commonly used words from the top 60 are indicated in italics in Appendix 1.

Table 5. Text Coverage of the Top 60 TOEIC Vocabulary in Both the Sections

	Word family	Common words in both section/ Whole running words of the listening section
Rate of coverage with the Reading section	26	43.33%
N I T L	35	56.67%
Sum	60	100.00%

(excluding function words)

(5) Other findings

Since most of the words that appeared in the top 60 word lists were 1000 level words, high level words that frequently appeared in the TOEIC vocabulary were extracted for reference. Appendices 2 and 3 indicate the high frequent AWL and the high frequent modified NITL words, respectively. In this section, the top 10 words are referred to as 'high frequent' due to the limitations of this paper. In addition, business terms were studied to further understand the features of the top 10 high level words. Furthermore, the words that were commonly used in both the reading and listening sections were investigated. Consequently, there were two findings. First, while 5 business terms were identified in the AWL words, there were none in the modified NITL words. Second, 3 words from the AWL and the same number of NITL words were commonly used in both the sections. These words are equip, confer, and compute from the AWL and personnel, supervise, and executive from the NITL words (see Appendices 2 & 3).

Table 6. Features of the Sectional Top 10 Words

Word Level	Section	Business terms	[business terms/ 10words]%	commonly used words in L & R sections
AWL	L	5	50.00%	3
	R	5	50.00%	
NITL	L	0	0.00%	3
	R	0	0.00%	

5. Discussion

(1) About half the words were found to be common in the listening and the reading sections.

Words in the reading section were set as basewrd, and subsequently, the rate of coverage with the words in the listening section was calculated. As a result, 52.63% of the words were found to be the same (see Table 3). The same calculation was conducted to determine the rate of coverage between the top 60 listening words and the top 60 reading words. The results showed that 43.33% of each category shared the same vocabulary (see Table 5). These two results suggest that it would be advisable to commence studying from these words, which are used in both the sections.

In order to perceive the characteristics of the commonly used words in both the

sections, 26 of the top 60 words were selected. Of these 26 words, 3 words (i.e. company, use, meet) were business terms and the others words were associated with time (i.e. year, month, week, day, hour, time), titles (i.e. Mr., Mrs.), and basic verbs (i.e. make, take, call, give, need). This shows that the high frequency words of the TOEIC vocabulary that were common to both the sections were also frequently used in everyday life. However, since most of the high frequent words were low level words or the first 1000 level words, the high frequent and higher level words were then selected (see Appendices 2 & 3). In Appendices 2 and 3, the words indicated in italics are the words to be prioritized for study because they are the words common to both the sections. These words are equip, confer, compute, personnel, supervise, and execute.

(2) The words that appear in the reading section are more academic and of a higher word level than those in the listening section.

AWL lists words that commonly appear in different kinds of academic texts. The findings of the study indicate that more words in the reading section can be classified under AWL than those in the listening section. From the modified NITL, we can discern the rate of low frequency words. Moreover, in the case of low frequency words, more words from the reading section can be classified under this category than those from the listening section. This indicates that—as could possibly be deduced from the beginning—we can conclude that the words used in the reading section are of a higher word level than those in the listening section.

It is generally assumed that the 2000 most frequent words will cover 80% of the text from any subject area (Mochizuki, Aizawa, & Tono 29). In the listening section, 80% of the words can be covered by a 2000 word level vocabulary, deducting the AWL and modified NITL words from the running words. However, in the reading section, a 2000 word level vocabulary can cover only 67.81% of all the running words. In order to tackle the reading section of the TOEIC, it is evident that we need to intentionally learn low frequency vocabulary in order to cover the remaining 12.19% of vocabulary that makes up 80% coverage of words. For this, the words listed in appendices 2 and 3 can be used for reference.

(3) In the reading section, business terms are frequently used, even in the top 60 most frequent words.

In general, it could be assumed that high frequency words are those used in everyday conversation and have lower word levels. In fact, 1000 level words comprised 96.67% of the top 60 words from the listening section, which also listed words that were often

used in everyday life such as go, here, and know.

On the other hand, in the reading section, 1000 level words constituted 85.00% of the top 60 words. However, as compared to the listening section, the reading section contained more business terms, which comprised 25.00% (15 words) of the top 60 reading section words. The 15 words included polysemous words (e.g. account, provide, use) and words with numerous derivative forms (e.g. develop, product, use). These also included words that could be used in a computer related context such as account, provide, and use. In the present study, 'computer related terms' were not included in business terms because that category did not appear in the 'TOEIC vocabulary list 1,2,3'. However, if this subject area of words is to be included, we could add words such as compute, line, and register to the business terms, and such an addition will raise the rate of coverage of business words in the reading section to 30%.

The top 60 words account for 11.80% of all the words in the mini-TOEIC corpus that was created for this study. Particularly in the reading section, being unacquainted with business terms means that one cannot cover 2.74% of the running words. This implies that it is necessary for learners to prioritize their study of the business related terms that appear in the top 60 word list.

(4) Words used in the listening section are of lower word levels than those in the reading section.

In the listening section, two of the four parts are conducted in an interactive conversational style and the remaining two parts employ a descriptive explanatory style. Thus, colloquial expressions appear frequently throughout the listening section. This is likely to be one of the reasons why proper nouns are less prevalent in the listening section than in the reading section. Moreover, lower level words such as 1000 level words and category b NITL words are more prevalent in the listening section than in the reading section. On the other hand, higher level words such as AWL and modified NITL words constitute only 18.72% of the listening section. This demonstrates that the words in the listening section are from a lower word level than those in the reading section. However, this fact alone does not decide the difficulty level of the questions in the listening section. According to Nation's classification of vocabulary knowledge, it is obvious that the listening and the reading sections of the TOEIC test different aspects of vocabulary. Thus, we can say that although the level of the words used in each section is likely to be of some help in preparing for the TOEIC, it is necessary to keep in mind that this alone will not determine the difficulty of the actual TOEIC questions. Although the words that appear in the listening section are from a lower word level and the examinee

may be able to recognize their written forms, it would be necessary to prepare for the listening section questions by undergoing training to recognize the 'receptive knowledge' regarding 'what the word sounds like', as suggested by Nation (27).

Limitation and Pedagogical Implications

The main limitation stems from the size of the corpus used in this study. In the present study, data to compile TOEIC corpus was specified to those taken from the Official TOEIC Guide and Exercise Book, vols. 1 & 2 because these were the only two official books published by the TOEIC test makers when the original version of this paper was presented. However, it is possible to point out that this had limited the data to be dealt with in the present study. Further investigations are needed to determine whether the conclusion of this study can be extrapolated with increased data that makes a larger corpus.

In May, 2006, the test style of TOEIC was renewed. There is an Official TOEIC Exercise Book that will prepare the examinees for the new test, and another just coming out. For the present study, these two books were excluded from the data analysis from the point of view of the time limit and to avoid possible influence of the difference of the test format. For future research, it would be also beneficial for the test takers to find out TOEIC vocabulary used in the new version of TOEIC and to find out differences in the TOEIC vocabulary caused by the change in the test format. This could be conducted in comparison with the result gained from the present study.

As for the pedagogical implication, to deal with TOEIC listening section, we can recommend test takers to have intensive listening training on words which test takers are familiar with when they are written but not so familiar with when they are pronounced. An example of such words would be English loan words which Japanese learners of English recognize when they are spelled out or pronounced in a Japanese way, or polysemous words which are pronounced differently according to the part of speech.

Moreover, it would be advisable for the test takers to start building their basic vocabulary from 2000 word level of the GSL. This will be of benefit for both listening and reading sections. To prepare particularly with the reading section, learning words from lower level but occurs frequently in TOEIC reading (such as words in appendices 2 & 3) and business terms is needed.

Works Cited

Coxhead, Averil. "New Academic Word List." TESOL Quarterly, 34.2(2000): 213-238.

Nation, Paul. Learning Vocabulary in Another Language. Cambridge: Cambridge University Press, 2001. 27.

Nation, Paul, and Newton, Jonathan. "Teaching Vocabulary." Ed. J. Coady, Second Language Vocabulary Acquisition, Cambridge: Cambridge University Press, 1997: 238-254.

Tanimura, Midori. and Utiyama, Masao. "Reading Materials for Learning, TOEIC Vocabulary Based on Corpus Data". JACET Bulletin, 42(2006): 81-96.

West, Michael. General Service List. London: Longman, 1953.

相澤 一美・小野和彦、「英検と大学用専門書の英文に使用される語彙の頻度調査」、第6回英語辞書学ワークショップ「英語の辞書と語彙」口頭発表、東京電機大学、2005年12月10日。

石川慎一郎、「TOEICの語彙-その特徴とレベル」、『言語文化学会論集』、19(2002):201-214。大学英語教育学会基本語改訂委員会(編)、『大学英語教育学会基本語リスト(JACET8000)』、大学英語教育学会、2003。

「TOEIC テスト採用学校一覧」、国際ビジネスコミュニケーション協会、2005年

<http://www.toeic.or.jp/toeic/data/pdf/TOEIC_2005.pdf>

国際ビジネスコミュニケーション協会、『TOEIC テスト活用実態報告』、13(2005) : 20。

中條清美、「英語初級者向け「TOEIC 語彙1, 2」の選定とその効果」、『日本大学生産工学部研究報告 B』 36(2003) : 27-42。

中條清美他、「ビジュアルベーシックによる TOEIC 用語語彙養成ソフトウェアの試作」、『日本大学生産工学部研究報告 B』 37(2004) : 29-43。

国際ビジネスコミュニケーション協会 (編)、『TOEIC 公式ガイド&問題集』、東京: 株式会社国際コミュニケーションズ、2000年。

国際ビジネスコミュニケーション協会 (編)、『TOEIC 公式ガイド&問題集 Vol.2』、東京: 株式会社国際コミュニケーションズ、2002年。

望月正道・相澤一美・投野由紀夫、『英語語彙の指導マニュアル』東京: 大修館書店、2003年。

「各種統計情報(平成17年度)」、文部科学省、2005年、

<http://www.mext.go.jp/b_menu/toukei/001/05122201/002/001/001.htm>。

Notes:

1 Range 29b is downloadable from the following site:

<<http://www.vuw.ac.nz/lals/staff/paul-nation/nation.aspx>>

2 This essay is based on the presentation made at the 1st Conference on Researching, Learning & Teaching Second Language Vocabulary of JACET English Vocabulary Research Group, 2005. Thus, the data used here are the most updated data of the time.

Appendices

Appendix 1. Top 60 High Frequent Words of TOEIC Vocabulary

rank	Listening	Word levels	Business term	Frequency of occurrence	Rate of frequency	Reading	Word levels	Business term	Frequency of occurrence	Rate of frequency
1	WORK	1		50	0.39	YEAR	1		69	0.47
2	OFFICE	1		48	0.38	EMPLOY	1	•	58	0.40
3	NEW	1		47	0.37	WORK	1		54	0.37
4	YES	1		46	0.36	OFFICE	1		53	0.36
5	GO	1		45	0.35	COMPANY	1	•	52	0.36
6	TAKE	1		42	0.33	NEW	1		48	0.33
7	MRS	1		39	0.31	MR	1		47	0.32
8	TIME	1		37	0.29	COURSE	1		41	0.28
9	WEEK	1		37	0.29	DAY	1		41	0.28
10	YEAR	1		36	0.28	INFORM	2		39	0.27
11	MAKE	1		32	0.25	TIME	1		38	0.26
12	SPEAK	1		32	0.25	BUSINESS	1		34	0.23
13	MEET	1	•	31	0.24	USE	1	•	33	0.23
14	THINK	1		31	0.24	MAKE	1		32	0.22
15	HERE	1		29	0.23	PAY	1		32	0.22
16	MAN	1		29	0.23	HOURLY	1		30	0.21
17	MR	1		29	0.23	PLEASE	1		30	0.21
18	BUILD	1		28	0.22	CUSTOMER	2		30	0.21
19	HOURLY	1		28	0.22	MEET	1	•	29	0.20
20	CALL	1		26	0.20	SALE	1		29	0.20
21	MONTH	1		25	0.20	SYSTEM	1		26	0.18
22	MORNING	1		25	0.20	TAKE	1		26	0.18
23	VERY	1		25	0.20	MRS	1		25	0.17
24	JUST	1		24	0.19	PLAN	1		25	0.17
25	ROOM	1		24	0.19	PRODUCT	1	•	25	0.17
26	BUSINESS	1		23	0.18	RECEIVE	1		25	0.17
27	COMPANY	1	•	23	0.18	WEEK	1		25	0.17
28	KNOW	1		23	0.18	TELEPHONE	2		25	0.17
29	NEED	1		23	0.18	CHARGE	1		24	0.16
30	WOMAN	1		23	0.18	DEPARTMENT	1	•	24	0.16
31	SO	1		22	0.17	INCREASE	1		24	0.16
32	TALK	1		22	0.17	INTEREST	1		24	0.16

33	TOO	1		22	0.17	MONTH	1		24	0.16
34	GOOD	1		21	0.16	ORDER	1		24	0.16
35	LONG	1		21	0.16	SERVICE	1	●	24	0.16
36	SALE	1		21	0.16	CALL	1		23	0.16
37	DAY	1		20	0.16	CHANGE	1		23	0.16
38	PEOPLE	1		20	0.16	SEND	1		23	0.16
39	PLACE	1		20	0.16	BUILD	1		22	0.15
40	PLEASE	1		20	0.16	PERSON	1		22	0.15
41	SEE	1		20	0.16	REQUEST	2	●	22	0.15
42	TODAY	1		20	0.16	COMPUTE	3		22	0.15
43	FLY	1		19	0.15	DATE	1		21	0.14
44	WAIT	1		19	0.15	NEED	1		21	0.14
45	SCHEDULE	3	●	19	0.15	NOTICE	1		21	0.14
46	BUY	1		18	0.14	NUMBER	1		21	0.14
47	CLOSE	1	●	18	0.14	MANAGE	2	●	21	0.14
48	HOUSE	1		18	0.14	PM	NITL		20	0.14
49	LEAVE	1		18	0.14	GIVE	1		20	0.14
50	RIGHT	1		18	0.14	TRAVEL	1		20	0.14
51	USE	1	●	18	0.14	REGISTER	3		20	0.14
52	TELEPHONE	2		18	0.14	ACCOUNT	1	●	19	0.13
53	BEGIN	1		17	0.13	FORM	1	●	19	0.13
54	LOOK	1		17	0.13	OFFER	1	●	19	0.13
55	NOW	1		17	0.13	PROVIDE	1	●	19	0.13
56	PLAN	1		17	0.13	SUPPLY	1	●	19	0.13
57	PLAY	1		17	0.13	CONTACT	3		19	0.13
58	WELL	1		17	0.13	DEVELOP	1	●	18	0.12
59	CHANGE	1		16	0.13	GAS	1		18	0.12
60	GIVE	1		16	0.13	LINE	1		18	0.12

* Word level '1' shows the first 1000 level words, '2' is for 2000 and '3' is for AWL level words.

Appendix 2. Top 10 of High Frequent AWL Words

rank	Listening	Business terms	Frequency of appearance	rate[%]	Reading	Business terms	Frequency of appearance	rate[%]
1	SCHEDULE	●	19	0.15	COMPUTE		22	0.15
2	JOB		15	0.12	REGISTER		20	0.14
3	AREA	●	12	0.09	CONTACT		19	0.13
4	COMPUTE		11	0.09	CONFER	●	18	0.12
5	CONFER	●	10	0.08	FUND		18	0.12
6	DESIGN		9	0.07	EQUIP	●	17	0.12
7	CONSTRUCT		8	0.06	FACILITATE	●	17	0.12
8	CONTRACT		7	0.05	INVEST		17	0.12
9	DOCUMENT	●	7	0.05	REQUIRE	●	17	0.12
10	EQUIP	●	7	0.05	AVAILABLE	●	16	0.11

Appendix 3. Top 10 of High Frequent NITL Words

rank	Listening	Business terms	Frequency of appearance	rate[%]	Reading	Business terms	Frequency of appearance	rate[%]
1	PERSONNEL		8	0.06	PERSONNEL		13	0.09
2	CELEBRATE		7	0.05	DEPOSIT		12	0.08
3	TRAFFIC		7	0.05	REFUND		11	0.08
4	CLIENT		6	0.05	SUPERVISE		11	0.08
5	SUPERVISE		6	0.05	EXECUTIVE		10	0.07
6	EXECUTIVE		5	0.04	REIMBURSE		9	0.06
7	PHARMACY		5	0.04	AUDIT		8	0.05
8	DENTAL		4	0.03	IMPORT		8	0.05
9	DEPARTURE		4	0.03	CERTIFICATE		7	0.05
10	FORECAST		4	0.03	COMPETE		7	0.05

要旨

TOEIC 使用語彙のセクション別比較分析

中山夏恵

近年、大学における英語教育はコミュニケーション能力を養成することを重視する風潮にあり、それを測定する指標として、TOEIC テストへの注目が急増している。この TOEIC の特徴の 1 つに、使用されている語彙数が多いことが挙げられる。また、その TOEIC の構成語彙であるが、「ある程度偏っていることを受験者たちは体験的に知っている(石川 201-214)」。これらの理由から、TOEIC テストの受験対策を行う上で、使用されている語彙、及びその特徴を明らかにすることには意義があるといえよう。

既存の単語集を見ると、TOEIC 使用語彙をセクション別、つまり、リスニングセクションに出題されているか、リーディングセクションに出題されているかという分類で扱っているものよりも、テストを総合的に見て出題頻度順に分類しているものが主流といえるであろう。Nation (27) は、語彙知識を計 18 の下位項目から成ると説明している。その分類に従うとリスニングセクションで試されている知識が、語形の下位項目の受容的音声に区分できるのに対し、リーディングセクションで試されている知識は、語形の下位項目の受容的綴りに区分出来るであろう。つまり、TOEIC においては、セクションにより、語彙の異なる様相を測っていることが指摘できる。もし、セクションにより、語彙の異なる様相を測っているならば、リスニング・リーディング各々のセクションでは出現する語彙も異なることが考えられる。しかし、セクション別に使用されている語彙、及びその特徴を探っている先行研究は、筆者の知る限りない。

本研究では、TOEIC 使用語彙をセクション別に調査し、その結果を比較することにより使用語彙におけるセクション別の特徴を明らかにすることを目的とする。セクション別の特徴が分かることで、よりの的を絞った TOEIC の受験指導が期待できるものと思われる。