

岩崎研究会創立 60 周年記念号

The Effect of Dictionary/App Usages in M/C Vocabulary Task	TOSHIKO KOYAMA TOMOKO YABUKOSHI	1
A Corpus-based Study on the Use of Descriptive Language by Japanese EFL Learners in Spoken Picture Description Tasks	RISA TERADA	17
An Analysis of Association Measures in Collocation Extraction from a Pedagogical Perspective	KOHEI FUKUDA	43
投稿規定		65
編集後記		65

Iwasaki Linguistic Circle

c/o Eigo-senko kyoiku junbishitsu
Tokyo University of Foreign Studies
3-11-1 Asahi-cho, Fuchu-shi
Tokyo 183-8534
Japan

Shoji Miyai, Chairman

2022年6月20日印刷 2022年6月29日発行

編集兼発行者 岩崎研究会 代表者 宮井捷二
振替口座 00140-2-8737

〒183-8534 東京都府中市朝日町 3-11-1 東京外国語大学 英語専攻教育準備室
email: iwaken.info@gmail.com

印刷所 有限会社紫藤印刷

The Effect of Dictionary/App Usages in M/C Vocabulary Task

TOSHIKO KOYAMA
TOMOKO YABUKOSHI

1. Background

Together with the evolution of technology, dictionaries continue to evolve. Hartmann pointed out by quoting the proposition concerning the ‘communicative shift’ theory in McArthur, who has posited in 1986 as follows:

...four major stages in the development of human interaction, . . . , each of which is associated with a different reference technology, from ‘oral’ and ‘script-based’ to ‘printed’ and ‘electronic’ dictionaries, with numerous subshifts and subtypes.

(Hartmann 41)

While the time migrates from analog to digital, the forms of dictionaries have gradually changed. The advent of pocket electronic dictionaries was a kind of an epoch-making event, as was described in detail by Sekiyama, especially in Japan. Koyama and Takeuchi summarized its popularity, and examined the relationship between college students’ lookup behavior and the retention of looked-up words or reading comprehension of English passages in using printed and pocket electronic dictionaries. In their series of studies, they found that pocket electronic dictionaries not only promoted learners’ look-up frequency more than printed dictionaries did, but also could reduce the time for L2 (Second/Foreign Language) reading. In spite of these advantages, however, it appears that this higher look-up frequency does not necessarily guarantee better reading comprehension nor retention of looked-up words.

While its popularity has been growing, Tono has already predicted in 2009 that “The future of pocket electronic dictionaries would be more integrative in nature. There will be a fuzzier boundary between PDAs, palm-top PCs, mobile phones and pocket electronic dictionaries.” (65). And his prediction has turned out somewhat to be true. Hubert reported that Japanese university students were switching to smartphone use as their primary dictionary resource from pocket electronic dictionaries. Collins’s description also supported this finding.

In line with the tendency, Koyama and Yabukoshi conducted a pilot study to explore Japanese college students’ use of gadgets and apps when they need to access lexical information in EFL classes, and examine the relationship with test scores. They especially focused on multiple-choice vocabulary quiz to perform their experiment. In the study, they found 1) most of students utilized a free apps with their smartphones; 2) pocket electronic dictionary users looked up more words than the users of smartphone apps; and 3) there were no statistically significant differences in English test scores in term of dictionary types. Their attempt, however, had some limitations such as the relatively small number of participants included in each group.

2. The study

2.1 Purposes

The current study aimed to investigate Japanese university EFL students’ dictionary use in a decoding task or a multiple-choice vocabulary task. In order to replicate the pilot study and examine if there are any differences in dictionary use over the two years, our findings will be discussed in comparison with Koyama and Yabukoshi, which had been conducted a year before the present study. We address the following two research questions:

RQ 1 What types of devices and dictionary apps are used by Japanese university students to look up unknown words in a multiple-choice vocabulary task?

RQ 2 Are there any differences in: (a) look-up behavior (i.e., the num-

ber of lookups and the time spent on the task); (b) learning outcomes; and (c) English proficiency levels in terms of students' dictionary choices?

2.2 Participants

A total of 73 college students (aged 18–19) participated in the study. They majored in health and sports sciences and were enrolled in a compulsory English reading and writing course at a university in the western part of Japan.

Table 1 The results of cloze test scores¹⁾

	<i>N</i>	<i>M</i>	<i>SD</i>
The present study	73	18.96	4.55
Koyama and Yabukoshi	97 ²⁾	18.10	4.94

Their English proficiency levels ranged from beginner to intermediate, which was similar to those in Koyama and Yabukoshi (Table 1). The results of *t*-test showed that there were no significant differences in the cloze test scores between the participants of the two studies ($t(168) = 1.16, p = .25, d = .18$).

2.3 Procedure

At the beginning of the semester, the participants took part in the experiment that was carried out in the same manner in Koyama and Yabukoshi. They were asked to answer 15 multiple-choice vocabulary questions, which were retrieved from the Part 5 of an official TOEIC[®] workbook. These materials seemed to include several unknown or unfamiliar words to the participants. Then, during the task, we allowed the participants to use their mobile devices, such as smart-phone apps and pocket electronic dictionaries, to look up unknown words where necessary. They were instructed to circle the looked-up words on the task sheet. After completing the task, they were asked to specify the types of mobile devices, apps, and/or dictionaries they had used. There were no time constraints imposed during the session. Additionally, the supplemental background survey was administered

to the participants to obtain information on their usual dictionary use, rather than the specific dictionary use at the time of the current experiment. The survey included questions as to the types of devices students possessed, dictionary apps they installed, online-dictionaries they accessed, and so on.

2.4 Data analyses

Analyses were conducted to answer the two research questions. First, students' responses to the question as to what mobile devices and dictionary apps they had used during the vocabulary task were analyzed. Then, the results of the present study and those in Koyama and Yabukoshi were compared to examine if there were any differences in preference of dictionary tools between the two years. To address the second research question, the participants were divided into four groups (i.e., three major dictionary groups and no dictionary group which did not utilize any dictionaries) based on the findings for the first research question. Regarding the look-up frequency, the number of words circled on the task sheet by the students of each group was counted. Then, due to the small sample size and the inequality of each group's sample size, Kruskal Wallis test, a non-parametric test, was conducted to examine if there were significant differences in the number of lookups among the three dictionary groups. As for the time to complete the task and the English test scores, Kruskal Wallis tests were performed to examine if there were significant differences in the time spent on the task and English test scores among the four groups. If significant differences were found by a Kruskal Wallis test, a post-hoc test (Mann-Whitney U test) was run to closely examine which group's mean was significantly different from each other. The results of the present study were, then compared with those in Koyama and Yabukoshi.

3. Results

3.1 Devices and dictionary apps

The current study found that most of the participants (80.8%) of the

study utilized smartphone dictionary apps and a handful of them (9.6%) used pocket electronic dictionaries to look up unknown words in the vocabulary task. As shown in Table 2, the use of smartphone dictionary apps has become more popular while pocket electronic dictionaries less popular, compared with Koyama and Yabukoshi. According to the supplemental background survey, all participants of the current study have their own smartphones. The survey also revealed that most of the participants (82.2%) also possess their own pocket electronic dictionaries. Based on the data we collected, the authors presume that students might not bring their pocket electronic dictionaries to the EFL classroom or prefer smartphones to pocket electronic dictionaries when looking up unknown words. The present study also revealed that 9.6% of the students did not rely on any dictionaries to complete the vocabulary task. That proportion has increased by 4.5% from the previous year.

Table 2 Comparison of the number and percentage of mobile devices

Devices	The present study		Koyama & Yabukoshi	
	<i>n</i>	%	<i>n</i>	%
Smartphone dictionary apps	59	80.8	74	75.5
Pocket electronic dictionaries	7 ³⁾	9.6	18 ⁴⁾	18.4
Unspecified	0	0	1	1.0
No dictionaries	7	9.6	5	5.1
Total	73	100	98	100

Regarding the types of dictionary apps, the results showed that the students utilized various smartphone dictionary apps. Weblio and Google Translate were the top two dictionary apps (Table 3). Weblio, free online 563 dictionaries with encyclopedia provided by GRAS Group Corporation, currently includes more than 9,860,000 entries in both English-Japanese and Japanese-English dictionaries. Weblio users can perform a bulk search on such multiple dictionaries, obtaining information about word definitions, pronunciations, and examples. Google Translate is a free translation service offered by Google. It

provides word-, phrase-, and sentence-level translations as well as its pronunciation guidance for each translation. Both Weblio and Google Translate have introduced a website interface and a mobile app for iOS and Android, so both services are available with a mobile phone in either online or offline mode. According to the background survey, most of the participants of this study had ever accessed the websites (79.5%) to search the meaning of a target word, and 17.8% of them had installed such dictionary apps in their smartphones. As for pocket electronic dictionaries, *Genius English-Japanese Dictionary*, which is one of the best-selling dictionaries among EFL learners in Japan, was commonly used by the pocket electronic dictionary users. Comparing the results of the present study and Koyama and Yabukoshi (Table 3), similar dictionary apps were used by the students in both studies. The following sections report on the use of the top three dictionaries (i.e., Weblio, Google Translate, and pocket electronic dictionaries) and its relationship with English test scores.

Table 3 Comparison of the number of dictionary apps used

		The present study	Koyama & Yabukoshi
Devices	Dictionary apps ⁵⁾	<i>n</i>	<i>n</i>
Smartphones	Weblio	35	34
	Google Translate	29	31
	Google	4	5
	LINE	1	4
	Yahoo	0	3
	ALC Ejirō	3	3
	Others ⁶⁾	1	7
Pocket electronic dictionaries	<i>Genius English-Japanese Dictionary</i>	7	15
	Others ⁷⁾	0	4

3.2 The number of lookups

This section reports on the look-up frequency of the top three dictionary users (i.e., Weblio, Google Translate, and pocket electronic

dictionary users). Those who reported using both pocket electronic dictionaries and smartphone dictionary apps were included in the pocket electronic dictionary group due to their minimum use of smartphone devices according to their English instructor's observation. Those who reported using both Weblio and Google Translate were excluded from the data analysis. Table 4 shows the number of lookups by the three dictionary groups of the present study and Koyama and Yabukoshi. A Kruskal Wallis test revealed that there were significant differences in the number of lookups among the three groups of this study ($H(2) = 14.07, p = .001$). The post-hoc test (Mann-Whitney U test) showed that the pocket electronic dictionary users looked up more words than the Google Translate users ($U = 9.00, p < .001, r = .65$) and the Weblio users ($U = 19.00, p = .001, r = .55$). There was no significant difference in the look-up frequency between the Google Translate and Weblio users ($U = 218.50, p = .176, r = .20$). These results support the findings in Koyama and Yabukoshi that there were significant differences in the number of lookups among the three dictionary groups and, in particular, that the pocket electronic dictionary group significantly looked up more words than the Google Translate group.

Table 4 Comparison of the number of lookups

Group	<i>n</i>	The present study			Koyama & Yabukoshi			
		The number of lookups			The number of lookups			
		<i>M</i>	<i>SD</i>	<i>Median</i>	<i>n</i>	<i>M</i>	<i>SD</i>	<i>Median</i>
Weblio	27	14.41	11.77	11.00	27	22.00	14.37	10.00
Google Translate	21	11.05	10.87	9.00	16	16.13	22.49	10.00
Pocket Electronic Dictionary	7	33.57**	10.94	35.00	18	32.83*	22.52	30.00
All	55	15.56	13.21	12.00	61	23.66	20.03	16.00

* $p < .05$, ** $p < .01$

3.3 The time to complete the task

Another look-up behavior that was investigated in this study was the time to complete the vocabulary task, which was not examined in

Koyama and Yabukoshi. The participants were divided into four groups including the three major dictionary groups (Weblio, Google Translate, and pocket electronic dictionary groups) and one group that did not use any dictionaries (No Dictionary group). As shown in Table 5, the pocket electronic dictionary group took a little longer time than the other three groups, and the No Dictionary group completed the task a little sooner than the other three dictionary groups. However, no statistical differences were confirmed in the time to complete the task among the four groups ($H(3) = 4.53, p = .21$).

Table 5 Comparison of the time to complete the task

Group	Time to complete the task			
	<i>n</i>	<i>M</i>	<i>SD</i>	<i>Median</i>
Weblio	27	28.41	6.08	29.00
Google Translate	21	29.14	7.15	28.00
Pocket Electronic Dictionary	7	33.86	8.82	34.00
No Dictionary	7	24.43	6.93	25.00
All	62	28.82	7.08	28.50

All values are n.s.

3.4 Learning outcomes

The four dictionary groups' learning outcomes were measured by the vocabulary task. Results of this study showed that the pocket electronic dictionary group seemed to gain higher scores on the word task than the other three groups, and that the No Dictionary group marked a slightly lower scores than the other three dictionary groups (Table 6). However, the statistical analysis did not confirm significant differences in the vocabulary quiz scores among the four groups of the present study ($H(3) = 6.03, p = .11$). These findings are almost consistent with those in Koyama and Yabukoshi, suggesting that, despite the frequent lookups by the pocket electronic dictionary group, there were no major differences in the task performance in terms of students' dictionary choices.

Table 6 Comparison of the vocabulary task scores⁸⁾

Group	The present study				Koyama & Yabukoshi			
	n	Vocabulary task			n	Vocabulary task		
		M	SD	Median		M	SD	Median
Weblio	27	6.63	1.94	7.00	27	6.30	2.60	6.00
Google Translate	21	6.48	2.36	7.00	16	6.44	1.93	6.50
Pocket Electronic Dictionary	7	8.00	2.38	8.00	18	6.78	2.65	6.50
No Dictionary	7	5.29	1.60	5.00	NA ⁹⁾	NA ⁹⁾	NA ⁹⁾	NA ⁹⁾
All	62	6.58	2.16	7.00	61	6.48	2.43	6.00

All values are n.s.

3.5 English proficiency

The participants' English proficiency levels were assessed by the cloze test. In the present study, the Weblio group seemed to obtain higher scores on the cloze test than the other three groups (Table 7). The results of a Kruskal Wallis test revealed that there were no significant differences in the cloze test scores among the four groups of this study ($H(3) = 1.55, p = .67$). These findings were supported by those in Koyama and Yabukoshi. Based on the research evidence, there seems to be no relationship between students' dictionary choices and their English proficiency levels.

Table 7 Comparison of the cloze test scores¹⁾

Group	The present study				Koyama & Yabukoshi			
	n	Cloze test			n	Cloze test		
		M	SD	Median		M	SD	Median
Weblio	27	19.93	3.15	20.00	27	17.44	5.24	18.00
Google Translate	21	17.71	5.73	19.00	16	20.25	3.22	20.50
Pocket Electronic Dictionary	7	18.43	2.88	19.00	18	17.11	5.50	16.50
No Dictionary	7	17.57	5.65	20.00	NA ⁹⁾	NA ⁹⁾	NA ⁹⁾	NA ⁹⁾
All	62	18.74	4.48	19.00	61	18.08	4.97	19.00

All values are n.s.

4. Discussion

The aim of the present study was to replicate and examine the findings in Koyama and Yabukoshi, which had investigated Japanese university EFL students' dictionary use in a vocabulary task a year before the present study. The findings of the present study are discussed and compared with those in Koyama and Yabukoshi in order to see if students' dictionary use has changed across the two years.

RQ1. What types of devices and dictionary apps are used by Japanese university students to look up unknown words in a multiple-choice vocabulary task?

The current research demonstrated that most students (80.8% of the participants of the study) reported using free dictionary apps, such as Weblio and Google Translate, in order to look up unknown words in the vocabulary task. The proportion of smartphone dictionary users has slightly increased compared to that (75.5%) in Koyama and Yabukoshi (20). These results may reflect the high permeation of the mobile devices in the society (MIC) and people's expectation of free access to online dictionaries (Dziemianko 5). Similarly, based on their three-year survey, Koyama and Yamanishi indicated that using free online translation such as Google Translate has become increasingly popular than paid dictionary apps. These findings suggest that Japanese college students would simply and effortlessly utilize such free translation tools that they have already owned and used since they were in high school, rather than bothering to choose and download a paid specific dictionary app. Regarding the use of pocket electronic dictionaries, only a handful of the students of the present study (9.6%) brought and consulted them to complete the word task even though almost all the students possessed such dictionaries at home (82.2%). The proportion of pocket electronic dictionary users has declined compared to that (18.4%) in Koyama and Yabukoshi (20). Moreover, 9.6% of the students in this study completed the word task without access to any dictionaries. That proportion has slightly increased in comparison to that (5.1%) of the finding in Koyama and Yabukoshi (20). These results may reflect

the tendency that students are less likely to depend on authentic dictionaries in L2 learning than before. In sum, the present study found that: (1) the students commonly preferred to use free online dictionary apps; (2) fewer students consulted pocket electronic dictionaries; and (3) the proportion of no dictionary users slightly increased from the previous year.

RQ2. Are there any differences in: (a) look-up behavior (i.e., the number of lookups and the time spent on the task); (b) learning outcomes; and (c) English proficiency levels in terms of students' dictionary choices?

Regarding the number of lookups, we found that the pocket electronic dictionary group consulted dictionaries more frequently than the Weblio and Google Translate groups. This finding is similar to Koyama and Yabukoshi (21), confirming that look-up frequency significantly differs in terms of students' dictionary choices. Despite the higher look-up frequency by the pocket electronic dictionary users, however, there were no significant differences in the time to complete the vocabulary task, learning outcomes assessed by the task, and English proficiency levels measured by the cloze test among the four dictionary groups (i.e., the three dictionary groups and the No Dictionary group). As for the time to complete the task, it was somewhat surprising that the two smartphone groups and the No Dictionary group took relatively as long as the pocket electronic dictionary group to complete the word task even though there were significant differences in the number of lookups. Regarding the smartphone users, Dziemianko (11) argued that the presence of advertisements displayed on the online dictionaries distracted the dictionary users and prolonged their search time in a receptive task. Similar findings were reported by Koyama (60), who examined the impact of dictionary interface (i.e., a smartphone dictionary or a tablet one) on look-up behavior and retention of the looked-up words. She conducted the experiment using both a smartphone and a tablet equipped with the same authentic dictionary, and found that the students looked up more words in a shorter period of

time when using a tablet dictionary as compared with using a smart-phone. In light of these findings, it could be assumed that the dictionary interface would not relate directly to the time spent on a language task, rather affecting search time and the number of lookups. As for the No Dictionary group, even though this group did not use any dictionaries, they took roughly the same amount of time as the other dictionary groups. This might be because no dictionary users had to infer the meanings of unknown words based on their lexical, syntactic and background knowledge to complete the task instead of looking up unknown words in dictionaries.

With respect to the relationships between dictionary choices and English test scores, no significant differences were found in the vocabulary task and the cloze test scores among the four dictionary groups. These findings indicate that students' learning outcomes and English proficiency levels did not differ in terms of their dictionary choices, despite the higher lookups by the pocket electronic dictionary group. In other words, a larger number of lookups using pocket electronic dictionaries does not appear to ensure higher scores on the vocabulary task. These results are in line with those in Koyama and Yabukoshi (22–24) and also supported by Koyama (59–60), which found that the frequent lookups using a tablet dictionary did not result in better performance on vocabulary and reading comprehension tasks. The empirical evidence thus suggests that there seem to be no immediate connections between students' dictionary choices and L2 learning performance. As Koyama and Yabukoshi (24) argued, other than individuals' dictionary choices or look-up frequency, their reference skills, strategies for dictionary use might bear a close link to learning outcomes as suggested by several previous studies (Koyama and Takeuchi; Mavrommatidou et al.). Investigating EFL learners' pocket electronic dictionary use, Koyama and Takeuchi (140–141) revealed that successful learners were good dictionary users, employing several strategies (i.e., using example search or idiom search to find further information, and/or looking up in more than two dictionaries), in contrast to less successful learners. More recently, a large-scale survey by

Mavrommatidou et al. (400) demonstrated that experienced dictionary users reported a higher degree of dictionary strategy use (i.e., familiarity with different types of electronic dictionaries and the conditions of their use, strategies for lemmatization and acquaintance with dictionary convention, and navigation strategies) than less experienced counterparts. In light of these findings, students' reference skills should warrant further investigation to shed light on critical factors contributing to better dictionary use and learning performance in L2 tasks.

5. Conclusion

The present study was conducted to replicate and examine the findings in Koyama and Yabukoshi and to investigate Japanese university students' dictionary choices and use in a vocabulary task over two years. The findings of this study in comparison to those in Koyama and Yabukoshi study suggested that: (1) smartphone dictionary apps (i.e., Weblio and Google Translate) have remained popular, pocket electronic dictionaries have become less popular, and the proportion of no dictionary users has slightly increased over the two years; (2) look-up frequency significantly differed in terms of dictionary choices—the pocket electronic dictionary users looked up more words than the other smartphone dictionary users; but (3) there were no significant differences in (a) the time to complete the word task, (b) learning outcomes assessed by the task, and (c) English proficiency levels measured by the cloze test, in terms of dictionary choices. These results provide evidence that frequent lookups using pocket electronic dictionaries are not likely to ensure better performance in the vocabulary task.

The present study as well as Koyama and Yabukoshi, however, was limited in their scope as they examined only the number of lookups and the time spent on a task as students' look-up behavior. Follow-up studies will be needed to further explore individuals' look-up behavior by means of qualitative methods (i.e., interviews, think-aloud protocols, and detailed analyses of video data) to obtain insights into how learners utilize dictionaries, particularly smartphone dictionary apps.

This line of research is promising because these gadgets have been commonly used by Japanese college EFL learners, but their effective use in L2 learning has remained insufficiently explored.

NOTES

* This article is a revised version of the paper presented by the authors at the FLEAT 7, 2019 in Tokyo, Japan.

- 1) The maximum score is 45.
- 2) A total of 98 students (aged 18–19) participated in the previous study. One of them was absent from the first session of the course, and his/her cloze test score was not available.
- 3) Among seven pocket electronic dictionary users, five of them used both electronic and smartphone dictionaries.
- 4) Among eighteen pocket electronic dictionary users, four of them used both electronic and smartphone dictionaries.
- 5) Multiple answers were allowed.
- 6) Others include a word navigation app, a translation app, an unknown dictionary app and so on.
- 7) Others include *English-Japanese Dictionary for the General Reader*, *O-LEX English-Japanese Dictionary*, a thesaurus, and an unknown dictionary.
- 8) The maximum score is 15.
- 9) No Dictionary group was not included in Koyama and Yabukoshi.

ACKNOWLEDGMENTS

The current study is partially supported by the Grant-in-Aid for Scientific Research (C) No. 19K00777, awarded to the authors of this article in the fiscal years from 2019 to 2021. The authors would like to express their deep gratitude to Professor Emeritus, Thomas Robb of Kyoto Sangyo University for his insightful comments on the earlier version of this paper.

REFERENCES

- Collins, John B. “Changes in electronic dictionary usage patterns in the age of free online dictionaries: Implications for vocabulary acquisition” *APU Journal of Language Research* 1, 2016: 36–49. <https://en.apu.ac.jp/rcaps/uploads/fckeditor/publications/languageresearch/APLJ_Volume1.pdf>
- Dziemianko, Anna. “The role of online dictionary advertisements in language reception, production, and retention” *ReCALL* 31, 2019: 5–22.
- Educational Testing Service. 2016. READING TEST PART 5. In *TOEIC® test koushikimondai*shu: *Shinkeishikimondaitaiouhen* [*The official TOEIC® test workbook: For the new item types*], 48–50, 90–92. Tokyo: The Institute for International Business Communication.
- Hartmann, Reinhard. R. K. *Teaching and researching lexicography*. Harlow, UK: Pearson Education, 2001.
- Hubert, Russell. P. “The use of smartphones as English dictionaries by university students in Japan and China” *Acta Humanistica et Scientifica Universitatis Sangio Kyotiensis* 50, 2017: 367–383. <<http://hdl.handle.net/10965/00002481Hu>>

- Koyama, Toshiko. "What is the desirable dictionary interface to EFL learners?" *Proceedings of the WorldCALL 2018: CALLing all the CALLers Worldwide, Concepción, Chile*, 2019: 58–60. <<https://worldcall.webs.upv.es/wp-content/uploads/2020/03/Proceedings-in-one-volume.pdf>>
- Koyama, Toshiko, and Osamu Takeuchi. "Comparing electronic and printed dictionaries: How the difference affected EFL learning" *JACET Bulletin* 38, 2004: 33–46.
- "Does an assigned task result in better retention of words?: Two empirical studies on handheld electronic dictionaries" *Language Education and Technology* 42, 2005: 119–132.
- "Does look-up frequency help reading comprehension of EFL learners? Two empirical studies of electronic dictionaries" *CALICO Journal* 25, 2008: <<https://journals.equinoxpub.com/index.php/CALICO/article/view/23125>>
- "How effectively do good language learners use handheld electronic dictionaries: A qualitative approach" *Language Education and Technology* 46, 2009: 131–150.
- Koyama, Toshiko, and Tomoko Yabukoshi. "University L2 learners' dictionary use in multiple-choice vocabulary quiz: A pilot study" *The Bulletin of Education, Osaka Ohtani University* 45, 2019: 17–27.
- Koyama, Toshiko, and Hiroyuki Yamanishi. "Daigakuseino eigogakushuujishono riyou: Sannen-anno ankeitochousawomotoni [Use of English language learning dictionaries by university students: The three-year questionnaire survey]" Paper presented at *Joint Conference of JACET English Lexicography Research Group, SIG on Reading & Vocabulary Acquisition Research Group, Waseda University, Tokyo*. 2019
- Mavrommatidou, Stavroula, Zoe Gavriilidou, and Angelos Markos. "Development and validation of the strategy inventory for electronic dictionary use (S.I.E.D.U.)" *International Journal of Lexicography* 32, 2019: 393–410. <<https://doi.org/10.1093/ijl/ecz015>>
- Ministry of Internal Affairs and Communications (MIC). 2019. Chapter 3: Basic data on the ICT field. In *2019 white paper on information and communications in Japan*, ed. MIC, 32–58. <<https://www.soumu.go.jp/johotsusintokei/whitepaper/eng/WP2019/chapter-3.pdf>>. 13 July 2020.
- Sekiyama, Kenji. "Twenty-five years of history and prospects of handheld unabridged electronic dictionaries in Japan — Galapagosization's gains and losses" *LEXICON* 46, 2016: 15–26.
- Tono, Yukio. "Pocket electronic dictionaries in Japan: User perspectives" in H. Bergehloltz, S. Nielsen & S. Tarp (Eds), *Lexicography at a Crossroads*, 2009: 33–67.

A Corpus-based Study on the Use of Descriptive Language by Japanese EFL Learners in Spoken Picture Description Tasks

RISA TERADA

1. Introduction

The proficiency expectations regarding English education in Japan has been evolving rapidly in recent years making the target English-skill levels necessary for Japanese students more explicit than before. For example, the Ministry of Education, Culture, Sports, Science and Technology (MEXT) recommends that Japanese students should get Grades Pre-2 to 2 in EIKEN before they graduate from high school. It is desirable for learners to have production skills that are more logical, coherent, and cohesive speeches and compositions. This present study investigates the nature of language features necessary for logical, coherent, and cohesive utterances by comparing native speakers with L2 learners and comparing learners with different proficiency levels. This study aims to reveal language features and developmental processes of spoken monologues using picture description tasks produced by Japanese learners of English. The author hopes that the results of this research will contribute to set clearer goals that will offer guidance to improve Japanese learners' communicative skills more efficiently in the future.

This study consists of six chapters with Chapter 2 reviewing previous major studies in relevant fields and Chapter 3 explaining the corpus used in the study focusing especially on picture description tasks. This chapter also covers the research method and procedures in detail. Chapter 4 shows the results of the data analysis, and Chapter 5 discusses the implications of the study results. Finally, Chapter 6

summarizes the major findings and suggests some limitations, and issues to be investigated for future studies.

2. Review of related literature

2.1. Previous studies focused on cohesion and coherence

Many researchers have studied learner language in terms of discourse organization and this section introduces previous related studies and summarizes learner features of discourse organization. Khalil (1989) evaluated English essays written by Arab college students from aspects of cohesion and coherence. He counted the number of cohesive elements included in every t-unit and classified them into five categories: references, substitutions, ellipses, conjunctions, and lexical cohesion. Table 1 indicates that the college students used repetitions most frequently in their essays. Repetitions and collocations are the components of lexical cohesion. In addition, references and conjunctions were also used frequently; however, the proportions of substitutions and ellipses were quite low. In coherence evaluations, the composition which elaborated information with specific examples got a high score. The correlation between the number of cohesive features and coherence evaluations indicated that the correlation was very low. Based on Khalil's research, compositions which include many cohesive features are not necessarily coherent.

Miyasako (2000), Sawaya and Suzuki (2016) had similar results to

Table 1. Types of Cohesive Ties (Khalil,1989)

Type of tie	Total	Percentage
Reiteration	117	61.9
Conjunction	34	18.0
Reference	29	15.3
Collocation	7	3.7
Substitution	2	1.1
Ellipsis	0	0.0
Total	189	

Khalil (1989) with Japanese participants. In Miyasako (2000), a class was divided into two groups based on their proficiency levels and students wrote essays in a regular class and were evaluated holistically. Then, cohesive features were counted and classified into five categories. The results showed significant differences in frequencies between the two proficiency groups. His students frequently used personal pronouns (references) and same word repetitions (lexical cohesion) as the cohesive features. Determining factors of holistic composition scores were also investigated using a multiple regression analysis. Only in the low proficiency group, the number of cohesive features was indicated as one of the determining factors of composition scores. From these results, the number of cohesive features does not necessarily relate to holistic evaluations.

Sawaya and Suzuki (2016) had their students write compositions after instructional focus on logical coherence and cohesion. As with the prior mentioned studies, their compositions were rated holistically and classified in terms of cohesive features. Their results revealed that their students frequently used personal pronouns and definite articles of references, conjunctions, and same word repetitions of lexical cohesion. As the correlation between the number of cohesive features and holistic evaluations, only personal pronouns indicated a positive correlation. They stated that this positive correlation related to the topic of compositions rather than the frequent use of personal pronouns.

2.2. Learner features of discourse organization focused on metadiscourse markers

Kobayashi (2009) conducted discourse analysis comparing native speakers of English and Japanese learners of English. This study used the Japanese Component of International Corpus of Learner English (the ICLE-JP) as the data of Japanese learners. The native speakers of English were extracted from the Louvain Corpus of Native English Essays (the LOCNESS). He used the classification of metadiscourse markers (Hyland, 2005) because metadiscourse markers help receivers to interpret and evaluate information appropriately. Hyland's (2005)

list is composed of ten categories with their functions and examples of each category are summarized in Table 2. He conducted a discriminant analysis and raised four metadiscourse features that contribute to the discrimination between native speakers and Japanese learners. Japanese learners overused frame markers, boosters, and self-mentions, while they underused hedges when compared to native speakers. Frame markers have a role in indicating discourse order, stages, and purposes. Japanese learners especially overused sequential expressions like: last, lastly, next, second, then, and third. Boosters can emphasize

Table 2. The categories of Metadiscourse Markers (Kobayashi, 2009)

Category	Function	Examples
Interactive resources	Help to guide reader through the text	
Transitions (TRA)	Express semantic relation between main clauses	in addition/ but/ thus/ and
Frame markers (FRM)	Refer to discourse acts, sequences, or text stages	finally/ to conclude/ my purpose here is to
Endophoric markers (END)	Refer to information in other parts of the text	notes above/ see Fig/ in section 2
Evidentials (EVI)	Refer to source of information from other texts	according to X/ (Y, 1990)/ Z states
Code glosses (COD)	Help readers grasp functions of ideational material	namely/ e.g./ such as/ in other words
Interactional resources	Involve the reader in the argument	
Hedges (HED)	Without writer's full commitment to proposition	might/perhaps/ possible/ about
Boosters (BOO)	Emphasize force or writer's certainty in proposition	in fact/ definitely/ it is clear that
Attitude markers (ATM)	Express writer's attitude to proposition	unfortunately/ I agree/ surprisingly
Engagement markers (ENG)	Explicitly refer to or build relationship with reader	consider/ note that/ you can see that
Self-mentions (SEM)	Explicit reference to author(s)	I/ we/ my/ our

speakers or writers' certainty and the expression *I think* was heavily used in this category. Self-mentions indicate explicit references to authors within their discourse. The Japanese learners of English heavily used the first personal pronoun, I and we.

2.3. Summary of learner features of discourse organization

This section summarizes learner features of discourse organization based on the results of previous studies. Learner features are summarized in two ways in this study. The first way is that learners use logical expressions like anaphoric and endophoric references, and conjunctions frequently in their discourse. The second way is that learners seem to have difficulties expressing their opinions and feelings efficiently by using hedges, boosters, self-mentions, and the like.

Logical expressions are really important factors to make discourse more transparent and their importance has been emphasized repeatedly in language teaching. However, overuse does not lead to coherent discourse. In the field of cohesion and coherence, many studies concluded that there is no, or weak, correlation between the number of cohesive features and coherence or holistic scores of language production. Cohesive features are generally considered mere factors contributing to coherence (Khalil, 1989; Miyasako, 2000; Crossley & McNamara, 2010; Sawaya & Suzuki, 2016). In other words, the increase in the number of cohesive features used in discourse does not necessarily lead to better discourse organization, which is also true for expressions related to speakers or writers' opinions, attitudes, and feelings. Kobayashi (2009) reported heavy overuse of the first personal pronouns by Japanese learners of English. As he stated, self-mentions can display its rhetorical effects by using them unexpectedly within an objective context. Nevertheless, in the case of Japanese learners, rhetorical effects of self-mentions are lost because subjective expressions account for most of the discourse. Japanese learners also extensively used the phrase, *I think* and it was interfered because of Japanese language influence and was not used appropriately. The overuse of one element repeatedly makes receivers feel uncomfortable.

Kobayashi (2009, 2010) mentioned that one of the causes of this tendency is from classroom instructions about paragraph writing. Examples of logical and sequence expressions are often introduced to instruct writing organizations in Japan (Nakagawa, 2016; Mita & Shimoda, 2021). These sequence expressions are helpful for learners to some extent however, focusing on only the surface features and rhetorical techniques causes learners to overuse one element and narrow their expressive variations. Examining only the quantitative data about surface features does not seem to offer any interesting suggestions for future instructions.

Researchers have stated that elaborateness of information and content also greatly influences transparent discourse (Khalil, 1989; Sawaya & Suzuki, 2016). Therefore, this study focuses on the more concrete aspects of discourse to obtain more beneficial findings. This study attempts to understand learner features of discourse, focusing on specific language expressions and what they intended to express in discourse.

2.4. Research question

Many researchers have investigated learner language in terms of discourse organization and previous studies have shown that there are differences of discourse organization between Japanese learners of English and native speakers of English. When instruction only focuses on surface features of discourse, it makes learners overuse these features, leading to non-target-like discourse structures. It is also important that learners need to decide what content should be included in their discourse.

With these perspectives in mind, the author explores some features of the spoken monologue by Japanese learners of English, with a special focus on the information selection and organization in picture description tasks. This study works on one research question “What information do speakers include as they describe the picture?” In order to reveal the content choices by speakers, the range of this research is specific and narrow, however, this research can provide

useful information. The following chapters, explain the methodologies used, the results of the research, and a detailed discussion of the results.

3. Method

3.1. The corpus used in this study

Learners' spoken data was extracted from the National Institute of Information and Communication Technology Japanese Learner English Corpus (NICT JLE Corpus, henceforth). The NICT JLE Corpus is one of the largest spoken learner corpora. The corpus includes 1,281 files by Japanese learners and 167 of them are tagged for grammatical and lexical errors. It also provides 20 native speakers' speech performance as a sub-corpus. This corpus was constructed based on the results of the speaking test called the Standard Speaking Test (SST, henceforth), which was designed and developed, following the ACTFL Oral Proficiency Interview (OPI). NICT worked with ALC PRESS to transcribe all the 1,200 audio files and annotated each transcript for spoken characteristics. Unfortunately, it does not include audio data (Izumi et al, 2004).

It takes each examinee approximately 15 minutes to have an interview test in SST. The SST is comprised of five stages: i) warm-up conversation, ii) a picture description task, iii) a role play, iv) a storytelling task and v) wind-down conversation. Each stage is divided into two parts: a task part and a follow-up part. In the task part, interviewees are asked to do specified tasks and in the follow-up part, interviewers have a question-and-answer sessions about the topics covered in the task with interviewees. The spoken data were evaluated by trained raters and scored based on their performance of all stages. Test scores are then converted to the SST level 1 (the least proficient) to level 9 (the most proficient).

3.2. Main focus of the present study

This research mainly used the spoken data of the second stage of the SST, the picture description task. Recall that examinees were

required to describe the situation of a given picture so that the examiners could imagine the content of the picture. The SST prepared seven versions of pictures: classroom, electric shop, map, neighborhood, restaurant, room, and skiing. The picture given to each examinee was randomly chosen among these seven pictures.

In order to analyze the descriptions from the detailed content, this study selected data from only one of the pictures because the content created from the seven pictures were quite different from each other and as a result, hard to control in terms of vocabulary. I selected the picture of a restaurant scene for this study and in this picture the interviewee describes a scene from a luxurious restaurant where different groups of people are having dinner. In order to communicate the atmosphere of the restaurant, the interviewee needs to describe the furniture, equipment, or decorations seen in the picture. They also need to mention about the people's behavior in the restaurant. The selection of what information to describe about the scene is really important in order for others to understand clearly what is the situation in the picture.

3.3. File selection

The section of the picture description task was extracted from each interview file before analyzing and Table 3 shows the number of files used in the analysis.

Table 3. The Number of Files Used in The Content Analysis

SST	Native	9	8	7	6	5	4	3	2
File	5	3	11	15	15	45	45	30	15

Depending on the SST level, the total number of files available to each SST level was quite different. For the lower-intermediate level, the SST levels 3 to 6, include the majority of learners while the beginner and the intermediate levels had a smaller number of files. This present analysis included all the files when the picture description task was about the picture of a restaurant, which was for the SST levels 2, 7,

8, and 9, as well as for native speakers. For the file extraction of the SST levels 3 to 6, 15 to 45 samples from each level were randomly selected. Regarding the SST level 1, it included only one file that matches the selection criteria for this study. Therefore, the SST level 1 was excluded in this analysis.

3.4. Content analysis

Content analysis of descriptive data was conducted to answer the research question: “What information do speakers include as they describe the picture?” To this end, it is essential to develop annotation schemes for coding picture information. In developing annotation schemes, since the NICT JLE Corpus does not provide actual pictures used in the SST, it was necessary to guess what was in the picture by looking at the content from the examinees’ answers. The coding category was made based on actual descriptive data mainly of native speakers and the more proficient learners.

In this study, two kinds of coding schemes were developed with the first being focused on informational content in the picture and it has two main categories: PLACE and PERSON. The PLACE category refers to all the information in their responses other than people in the picture, while the PERSON category refers to the information about people in the picture. The second coding scheme focused on utterance functions: DESCRIPTION and INTERPRETATION. The functional classification coded as DESCRIPTION is related to the extent of verifiability and the information which was objectively identified from in the picture. On the other hand, the information adding their own interpretations based on the picture was coded as INTERPRETATION. Below are a few examples of the coding used.

(1) PLACE:

- a. And ur urr there’s a number of wine glasses on the table.
(file11-native)
- b. And the time is seven o’clock. (file00654-SST level 7)

(2) PERSON:

- a. Mmm a man and a woman are in a restaurant. (file01211-SST level 9)
- b. Ur sommelier's in a tuxedo. (file01243-SST level 9)
- (3) DESCRIPTION:
 - a. And further in the background, there is some trees. (file11-Native)
 - b. And er one woman is playing the piano. (file00654-SST level 7)
- (4) INTERPRETATION:
 - a. It looks like a French restaurant. (file01236-SST level 8)
 - b. And he said, "It's good". (file00059-SST level 6)

The coding unit was a sentence and identification of sentences was made based on the transcriptions of the NICT JLE Corpus, where each sentence was delimited by their <s> tags. Each sentence was coded from two aspects, whether the sentence referred to PLACE or PERSON and whether the sentence belongs to DESCRIPTION or INTERPRETATION. There were some cases where more than two types of content were included within a single sentence. Also, in some sentences connected a clause of DESCRIPTION with a clause of INTERPRETATION and in these cases, the sentence was coded separately. Complete coding examples are presented here. In (5c), the former half of the sentence, "So um he has his own drink", was judged as the DESCRIPTION while the latter half, "which is his usual but I assume he's gonna probably try the new one in a minute" was judged as the INTERPETATION. The coding process was done manually by the author.

- (5) Coding examples:
 - a. script: And er one woman is playing the piano. (file00654-SST level 7)
annotation: PERSON-DESCRIPTION
 - b. script: It looks like a French restaurant. (file01236-SST level 8)
annotation: PLACE-INTERPRETATION
 - c. script: So um he has his own drink, which is his usual but I

assume he's gonna probably try the new one in a minute.
(file01234-SST level 8)

annotation: PERSON-DESCRIPTION

annotation: PERSON-INTERPRETATION

3.5. Statistics used in this study: HCFA (Hierarchical Configurational Frequency Analysis)

The frequency data of the content analysis was further analyzed by a hierarchical configurational frequency analysis (HCFA) to determine what information is preferred or not preferred by the different proficient speakers as they describe the same picture. HCFA is an extended version of a chi-square test. HCFA can test the relationships of a multi-dimensional table while the chi-square test is usually applied to a two-dimensional table. This test can examine whether the observed frequencies of variable level combinations are significantly different from the expected frequencies that are expected by chance. The variable level combinations are called *configurations* and “If a configuration is more frequent than expected, it is referred to as a *type*; if it is less frequent than expected, it is referred to as an *antitype*” (Gries, 2009: 244). In this study, configurations were based on three variables, PLACE/PERSON, DESCRIPTION/INTERPRETATION and SST levels.

HCFA processing was done with R based on the R script and a manual introduced in Gries (2009). In processing, the author is aware that a chi-square test should not be performed on proportional data, however this study normalized the mean frequency data of one file against 30 files since the dataset for native speakers were too small. In addition, learner files were grouped together based on their SST levels and corresponding CEFR-J levels according to Tono (2013) in order to highlight features of each level more explicitly. SST levels were aggregated into four classes: native, B, A2 and A1. A breakdown was the following: native (native), B (SST levels from 6 to 9), A2 (SST levels 4 and 5), A1 (SST levels 2 and 3).

4. Results

4.1. Results of the content analysis

Table 4 presents the mean frequencies of the two main content categories: PLACE/PERSON. Mean frequencies were applied in this table because the sample size was quite different among the levels. It becomes easy to compare the values using mean frequencies. This table shows that the mean frequencies increased as speakers' proficiency levels improved with native speakers producing the highest value. Looking at each main category, in general, speakers referred to the PERSON category more than the PLACE. The proportional use of each category within a description is different when comparing learners and native speakers. In the learner group the proportion of the PERSON is almost twice as large as the PLACE. However, the native speakers referred to the PLACE more frequently than learners, and used PERSON and PLACE in almost equal amounts.

Table 4. The Mean Frequencies of Each Content

SST	PLACE	PERSON	ALL
Native	7	7.2	14.2
9	5.3	7.7	13
8	3.6	9.1	12.7
7	3.7	6.4	10.1
6	3.3	6.4	9.7
5	3.5	7	10.5
4	2.4	6.3	8.7
3	2	4.3	6.3
2	1.5	2.7	4.2
ALL	32.3	57.1	89.4

Table 5 summarizes the average frequencies of two main functional categories: DESCRIPTION and INTERPRETATION. From this table, the mean frequencies of the DESCRIPTION are higher than the INTERPRETATION and this is not surprising that speakers included verifiable information more often than their own interpretations. They were asked to describe the given picture during the SST.

As an interesting point, the intermediate level learners added their own interpretations more frequently than the other groups.

Table 5. The Mean Frequencies of Each Utterance Function

SST	DESCRIPTION	INTERPRETATION	ALL
native	11.2	3	14.2
9	10	3	13
8	8	4.7	12.7
7	6.7	3.5	10.2
6	5.7	4	9.7
5	6.5	4	10.5
4	6	2.8	8.8
3	4.8	1.5	6.3
2	3.2	1	4.2
ALL	62.1	27.5	89.6

4.2. Results of HCFA

In this section, the results of HCFA are presented and explanation of how to interpret tables precede the results. The “content” means the variable PLACE/PERSON, while the “function” refers to the variable DESCRIPTION/INTERPRETATION, and the “sst” means the SST levels, respectively. The column named “Freq” shows the observed frequencies of each variable or variable interaction, whereas the column named “Exp” shows the expected frequencies. “Cont. chisq” stands for a contribution to chi-square since the HCFA provides the results of a chi-square test for each subtable. The chi-square value of a whole table is calculated by adding up differences between observed frequencies and expected frequencies in all cells. From the contribution to chi-square, the breakdown of the chi-square value in a whole table can be represented. The column named “Obs-exp” indicates types or antitypes for each cell. If an observed frequency of a cell is larger than an expected frequency, the cell is regarded as a type and is expressed by using >. In contrast, if an observed frequency is smaller than an expected frequency, the cell is regarded as an antitype

and is expressed by using $<$. In this study, type cells were marked in dark gray and antitype cells were marked in light gray. The “P.adj. Holm” is the adjusted p-value, because in processing HCFA with R, we can choose two adjusted p-values for multiple post hoc tests: the Bonferroni correction, and the Holm correction. Gries (2009) recommended using the Holm correction because it guarantees that an overall probability of error does not exceed 0.05. These corrections are quite important so that the research does not reject null hypotheses by mistake. The “Dec” indicates the degree of significance of each adjusted p-value and the degree of significance is expressed at three stages with asterisks. If a cell is highly significant, the column provides three asterisks (***). However, if the significance of a cell is low, the column provides one asterisk (*). Finally, in the case where a cell is not significant, the column provides “ns”. The rightmost column named “Q” provides the size of the effect and the larger the value of Q, the stronger the configuration contributes to the overall results.

Table 6 indicates the results of the HCFA for the interaction between PLACE/PERSON x SST levels and there are three significant configurations in this table. The configuration of PLACE x native is reported as a type with high significance. The configurations of PERSON x native and PLACE x A2 are considered antitypes. All of the other configurations are not considered as significant. This table is

Table 6. The results of HCFA for the (PLACE/PERSON x SST levels)

content	sst	Freq	Exp	Cont.chisq	Obs-exp	P.adj.Holm	Dec	Q
person	native	324	392.6598	12.0057	$<$	0.0002506	***	0.049
place	native	315	246.3402	19.1368	$>$	2.49E-05	***	0.044
person	A2	297	265.4602	3.7473	$>$	0.1037473	ns	0.02
place	A2	135	166.5398	5.9731	$<$	0.0288289	*	0.019
person	B	319	298.0282	1.4758	$>$	0.0980049	ms	0.014
place	B	166	186.9718	2.3523	$<$	0.1652333	ns	0.013
person	A1	171	154.8518	1.684	$>$	0.1911403	ns	0.01
place	A1	81	97.1482	2.6842	$<$	0.1936447	ns	0.009

in agreement with the data in Table 4. Native speakers doing the picture description tended to include PLACE elements more frequently than did the learners. Moreover, they did not describe PERSON elements more frequently than did the learners.

Table 7 summarizes the results of the HCFA for the interaction between DESCRIPTION/INTERPRETATION x SST levels. The configurations of DESCRIPTION x native, INTERPRETATION x B, and INTERPRETATION x A2 are regarded as types with low significance. On the contrary, the configuration of INTERPRETATION x native is regarded as an antitype with high significance. This table can explain the findings presented in Table 5 because native speakers and beginners were more likely to describe verifiable information, while the intermediate learners were more likely to include their own interpretations in picture description.

Table 7. The results of HCFA for the (DESCRIPTION/INTERPETATION x SST levels)

function.	sst	Freq	Exp	Cont.chisq	Obs-exp	P.adj. Holm	Dec	Q
description	native	504	453.0962	5.7189	>	0.0203187	*	0.038
interpretation	native	135	185.9038	13.9384	<	0.0001883	***	0.031
description	B	310	343.8993	3.3416	<	0.0865532	ms	0.023
interpretation	B	175	141.1007	8.1443	>	0.0155996	*	0.02
description	A2	279	306.3186	2.4364	<	0.1352168	ns	0.018
interpretation	A2	153	125.6814	5.9381	>	0.0390396	*	0.016
description	A1	189	178.6858	0.5954	>	0.2183819	ns	0.006
interpretation	A1	63	73.3142	1.4511	<	0.238826	ns	0.006

5. Discussion

5.1. Summary of major findings

The research question is “What information do speakers include as they describe the picture?” With this question in mind, content analysis was carried out using picture description data in the NICT JLE Corpus. The data was annotated according to the two coding schemes:

one focusing on picture content, and the other examining utterance functions. The annotation schemes were developed in a bottom-up manner by considering test examinees' actual descriptive data. The coding scheme focusing on the picture content was divided into the PLACE and the PERSON, while the other coding scheme focusing on the utterance functions was divided into the DESCRIPTION and the INTERPRETATION. Each sentence was classified manually by the author and after coding the data, the frequencies were counted. In addition, the author conducted hierarchical configural frequency analysis (HCFA) to test the contribution of certain configurations to the overall frequency changes across the speakers' proficiency.

The results show the findings are related to the general features and the features specific to each group. First, the amount of information generated increased with speakers' proficiency levels. Second, speakers preferred describing the information related to the people in the image rather than the places, and they also tended to include verifiable information more than supplying their own interpretations.

A comparison between learners and native speakers revealed that the proportion of information on the PLACE and the PERSON was different. While native speakers described both categories almost equally, learners described the PERSON category much more often. In addition, native speakers and novice-beginners included more verifiable information, while low-intermediate and intermediate learners preferred adding their own interpretations in their descriptions.

5.2. Implications for differences between native speakers and Japanese learners

The major findings of the present study revealed that there are some interesting differences in terms of the information that is included in the picture description. One of the differences is related to the distributions of the PLACE and the PERSON categories. Native speakers described a wide range of information on both the PLACE and the PERSON and in particular, they described the information about the place very differently from learners. For example, in (6a), this native

speaker described the distances between each table and the appearance of the entrance of the restaurant in a very specific way in order to describe the atmosphere of the restaurant. In another example(6b), a different native speaker also mentioned the rope in the entrance perhaps showing the restaurant was exclusive. In addition to these specifics, native speakers frequently mentioned the people in the picture who were not located at the center of the picture. Overall, native speakers' descriptions covered more information, and each description seemed to serve a particular discourse function.

(6) Examples of the PLACE by native speakers:

- a. And there are only three table in this picture, and they look like they're really far apart. So maybe it helps people to have their own space and not have to listen to other people's conversation. There's also a velvet rope that er seems to be blocking one of the entrances. And it almost looks like the entrance there's there's no door, like it goes straight outside. So maybe this is a very warm place to to have a restaurant. (file1-native)
- b. Ur there's a doorway with er some kind of rope strung across it. (file11-native)

On the other hand, the learners paid more attention to the PERSON category only. They focused on detailed information, and in particular, about the people at the center of the picture. Examples describing the clothes of the guests are shown in (7). While native speakers use phrases just like *dressed up* and *formal dresses and suits*, the learners even mentioned the colors and accessories.

(7) Examples of the guest-clothes:

- a. Well it looks like a very fancy restaurant where the waiter is dressed up in a tuxedo and he's serving wine to the customers who are also very dressed up in er formal dresses and suits. (file1-native)

- b. And there are some other nicely dressed couples in the background. (file14-native)
- c. And man is wearing yellow suit. And the woman is wearing a pink dress with pearl necklace and earrings. (file01281-SST level 7)

When learners described the information about the place, they mainly referred to concrete furniture: *the clock*, *the table*, and *the piano*, among others. The restaurant was also frequently explained and most explanations were related to a type of restaurant, and seemed to be related to topic introduction. (8) shows examples which described the restaurant.

(8) Examples of the restaurant:

- a. It looks like a French restaurant. (file01236-SST level 8)
- b. Eh I think eh Western eh European type. (file00116-SST level 5)

As a possible reason for this difference, the way the space is recognized may be different between learners and native speakers. Native speakers seem to regard the information about the place just as important as the information about the person. Abstract spatial aspects like the entrance, distances of tables were described as well as specific pieces of furniture and this was a unique feature for native speakers. However, learners constructed the picture description mainly focusing on the information about the people. Therefore, they seem to consider the information about the place as merely the background and did not emphasize them. The feature that learners focused on the people more often than the places would be related to some findings in Izumi (2013). She stated that learners rarely used inanimate subject sentences, while native speakers often used them during the picture description task. In her discussion, Izumi postulated that the different restrictions on both English and Japanese verbs may have an influence on this tendency. Differences of language restrictions might also lead to the different

recognition of space.

Another interesting difference is the degree of speakers' interpretations within the descriptions. In this task, interviewees were asked to describe the content of the given picture. It is natural that more verifiable information in the picture was expressed compared to speakers' own interpretations. However, the findings revealed that the lower-intermediate and intermediate learners included their own interpretations more often than the native speakers. Rich interpretations were added especially to the focused area of the picture. For example, many sentences that were guessing the relationships between guests were found from these groups of learners. On the contrary, these explanations were rare in native speakers' descriptions. From examples in (9), we can see that while native speakers referred to the relationship of guests simply, learners tried to describe them in detail.

(9) Examples of the relationship of the guests:

- a. Erm in this picture, it looks like there's a man and a woman out on a date. (file14-native)
- b. So in the night, the couple, maybe they are not so familiar to each other. Because that if they are they already had they already develop the strong relationship each other, that they don't maybe they don't need to be in this kind of formal expensive restaurant. So I think that they are er very in a beginning of their relationship. (file00980-SST level 8)

It may be that the way speakers thought in their mind may be different between learners and native speakers. The Standard Speaking Test, which the NICT JLE Corpus was based on, does not make available the evaluative criteria of picture description tasks. Thus, each test candidate might work on the picture description task in their own way. Native speakers seemed to transmit the information in the picture description objectively and generally. Therefore, their descriptions for each piece of information were simple and did not include many personal interpretations. On the other hand, learners seemed to

try to express the information in the picture in more creative ways and use their imaginations more freely. In (10), one can see some examples which describe the picture by making original stories, which are unique to a group of Japanese learners of English.

(10) Examples of the original story:

- a. This woman is er Momoko. And this man is Kentaro. And they are a umm they're good friends and decided to er come to have dinner together. And er they often date, but this time, they decided to er come to a good restaurant because er Kentaro tried to er propose her. (file00042-SST level 6)
- b. Um Mr. and Mrs. Yamamoto and er went to the restaurant to have the to have dinner. And uh they when they were engagement, they used to go to the restaurant such kind of restaurant er very often, but now er they have the children, so she ah they can hardly go to the such kind of fine restaurant. And er that night, they could find the baby sitter and er they could come to the er to the restaurant to have dinner. (file00059-SST level 6)

5.3. Implications for differences between different proficiency learners

What are the differences between learners at the different proficiency levels? Possible reasons for distinguishing their scores are discussed and relevant factors may be different for novice-beginners and lower-intermediate learners.

First, novice-beginners are likely to miss the basic information in their descriptions. For instance, information about the restaurant, the clock, and the guests are the most characteristic information in the picture. If speakers can mention at least these items, then the listeners can have an image of the picture more easily. However, beginners cannot even explain the most basic information very well. It may be that the beginners lack the vocabulary to express even the most basic elements of the picture. Although some of them tried to mention some of

these items, the information was too limited for the listeners to understand the picture appropriately. From (11a), the learner probably tried to describe the guest and the waiter, however, this description lacked other specific information. Moreover, they might also lack even the basic language skills needed to complete the task. This tendency matches with results obtained from Izumi (2013). Her analysis also revealed that beginners have difficulties making accurate sentences and ordering them cohesively. (11b) shows the characteristic information about the picture, but each sentence included many grammatical errors and fillers. The sentences are also not connected to each other very well. These aspects may make their descriptions less intelligible to listeners. Therefore, they need to improve vocabulary, grammar, and fluency to express the basic information in a structured way, rather than broadening the coverage of additional information.

(11) Examples of novice-beginners:

- a. He put up glass in uh uh in wine wine in glass. And he talked with eeto nanchuuno eh with waitress waiter. So eeto maybe he asked waiter wine's wine's uh what kind wine. (file01139-SST level 2)
- b. Er there comes party. And she plays piano. Mmm. She ah he have a glass. Waiter have wine. This time is seven o'clock. They have dinner. (file01133-SST level 3)

The descriptions by lower-intermediate learners had fewer grammatical errors, fillers, and repetitions when compared to the novice-beginners. The amount of information also increased gradually. However, they tended to heavily focus on a limited area of the picture, as shown in (12). Although this learner can explain the center part of the image really specifically, the information about the other people and places were not mentioned. From this explanation, the listeners can only get a limited amount of information about the picture. Higher level learners were able to mention the information about the place as a whole, and then the people at the center table, as well as the other

parts. In order to improve their description, lower-intermediate learners would need to broaden the range of information that they provide.

(12) Examples of lower-intermediate learners:

I think mmm this two people urr is mm was was married from, ur mm I think, ten years ago. So urr ten ten years mm anniversary, ur to they come to the restaurant come to this restaurant. And urr I think they they're they they like drinking. So, now, they selected wine ur select wine. Mm. Which is a ur he asked him which is a best wine for me for us. Mm. So ur the mm shop's owner bar bartender ur ur recommend this wine to him.
(file00610-SST level 4)

5.4. Pedagogical implications

There are pedagogical implications that learners can understand to construct more intelligible descriptions based on the major findings and possible reasons for higher scores. There are three implications on future instructions. The first implication is that teachers should prepare and present task purposes explicitly when they carry out picture descriptions. This study revealed that learners and native speakers approached picture description tasks quite differently. One of the reasons for this difference may be related to how they perceived the purpose of the task. While native speakers seemed to describe the overall information in the picture with simple and objective expressions, Japanese learners seemed to describe about a focused area of the picture with more subjective feelings, impressions, and with objective expressions in detail. Appropriate expressions and organizations are different depending on task purposes and we cannot judge which style is more appropriate. Therefore, to define the task purposes in advance is an essential factor. If the purpose is to describe more objectively, then learners can participate in a brainstorming activity together to check the overall information in the picture. In contrast, if the purpose is making a story, then it would be good to discuss the settings of the people and the situation with their full imagination.

The second implication is that it is important for beginners to start practicing from small areas rather than using the entire picture. The findings show that beginners may lack the basic language ability like vocabulary and grammar to successfully accomplish the task. To focus on the various aspects of the task at the same time would be a difficult for them. In order to improve gradually, they would need to practice describing with only one or two sentences and about a limited area. Gradually, this would lead to further steps for constructing sustained descriptions and later adding their own interpretations.

The third implication is that it is necessary to introduce language activities that more proficient learners can do to improve the quality of their picture description. They will become more able to make sustained speech, fluently and cohesively, to some extent. Their problems are based in the amount of information and a limited repertoire of language expressions. Teachers should introduce further language activities so that they can broaden the learners' viewpoints and language variations. For example, learners can check their picture description with each other after they complete it. Through listening to others' descriptions, they can learn new perspectives on the information and different language expressions from each other.

6. Conclusion

6.1. Statements of the present study

This study examined the spoken monologue from various aspects in order to get beneficial suggestions for improving Japanese learners' speaking ability. It is necessary for them to speak their opinions more logically, coherently, and cohesively. To reveal language features among different proficient speakers, a research question, "What information do speakers include as they describe the pictures?" was posed in this study. Two analyses were carried out to answer the research question: content analysis and HCFA (hierarchical configural frequency analysis).

This study could provide interesting findings in terms of the spoken monologue and there were different tendencies between Japanese

learners of English and English native speakers, as well as differences between learners with different levels of proficiency in English. As a difference between learners and native speakers, while native speakers tended to cover a wide range of information with verifiable, simple, and objective expressions, learners tended to describe a limited range of information. Moreover, learners' description included many subjective interpretations. In general, more proficient learners could provide more information during the task. According to the major findings, in order to construct logical, coherent, and cohesive discourse, learners need to care about the informational choice and presentation of their information, using basic language ability like vocabulary and grammar.

6.2. Limitations

This section discusses limitations of the study and the first limitation relates to the coding process because coding the information was carried out only by the author. In fact, reliability and validity are really important factors for discourse analysis and content analysis. This present study should have found another annotator to keep inter-annotator reliability, and to ensure the validity of coding. The second limitation is about task purposes of the picture description in the NICT JLE Corpus. As stated in Chapter 5, the specific task purposes of the picture description task were not indicated explicitly. Although learners and native speakers included the information in the picture differently, we cannot judge which style is more appropriate. If explicit information about the task purposes were available, implications of differences could have been discussed from different perspectives. The last limitation is related to the scoring systems of the SST since the SST is composed of five stages and examinees are asked to finish the interview test in 15 minutes. The spoken data based on the SST was accompanied with the SST score 1–9, which was the total performances of all stages. Thus, the final SST level might not necessarily indicate the level of performance in the picture description task. This might skew the analysis of the corpus data classified by the overall SST levels.

6.3. Future perspectives

This section mentions future perspectives as conclusions to this study. This present study analyzed picture description data in terms of concrete aspects including the information used in the description. As a result, general features and interesting differences were found between L2 learners and native speakers as well as between learners with different proficiency levels. However, there are further opportunities to analyze the picture description tasks. In this study, the overall discourse structures could not be dealt with. Based on the results obtained from this study, general structural patterns of picture description could be extracted. Through applying techniques of move analysis, further interesting findings and implications may be revealed. Moreover, this study chose to focus on only one picture as a research target however the NICT JLE Corpus has an additional six pictures. Although the picture description tasks at Stage 2 uses a single picture, the storytelling tasks have interviewees tell stories using several of them. It would be possible to compare the information in the picture, language expressions and structural patterns depending on the themes as well as the numbers of the pictures. The present author hopes that this study has shed some light on the possibility of corpus approaches toward the study of discourse organization processes for descriptive tasks by L2 learners of English.

NOTE

This is a shortened version of my MA thesis submitted to the Graduate School of Global Studies Tokyo University of Foreign Studies in partial fulfillment of the requirements for the degree of MASTER OF ARTS in Linguistics.

REFERENCES

- Crossley, Scott A and McNamara Danielle S. "Cohesion, coherence, and expert evaluations of writing proficiency." *Proceedings of the Annual Meeting of the Cognitive Science Society* 32 (2010): 984–989. <https://www.researchgate.net/deref/https%3A%2F%2Fescholarship.org%2Fuc%2Fitem%2F6n5908qx>
- Gries, Stefan Th. *Statistics for linguistics with R*. Walter de Gruyter, 2009. <https://doi.org/10.1515/9783110216042>
- Hyland, Ken. *Metadiscourse: exploring interaction in writing*. Continuum, 2005.

- Izumi, Emi, Isahara Hhitoshi, and Uchimoto Kiyotaka. *Nippon-jin 1200-nin no eigo speaking corpus [A spoken corpus of 1200 Japanese EFL learners]*. ALC Press, 2004.
- Izumi, Emi. “Nippon-jin gakusyusya no eigo hatsuwa ni okeru ronritenkairyoku [Logical construction of English speaking by Japanese learners of English].” *Eigo gakusyusya corpus katsuyo handobukku [A handbook of learner corpus research]*. Ed. Tono, Yukio, Kaneko Asako, Sugiura Masatoshi and Izumi Emi. TaishukanShoten, 2013. 178–186.
- Khalil, Aziz. “A study of cohesion and coherence in Arab EFL college students’ Writing.” *System* 17, 3 (1989): 359–371. [https://doi.org/10.1016/0346-251X\(89\)90008-0](https://doi.org/10.1016/0346-251X(89)90008-0)
- Kobayashi, Yuichiro. “NS/NNS tekisutobunruimoderu ni motozuku nippon-jin eisakubun no tokucyocysyutsu [Extracting characteristics of English essays written by Japanese EFL learners on the NS/NNS text categorization model].” *Jinmonkon 2009 Ronbunshu [Jinmonkon collections of academic papers 2009]* 16 (2009): 261–268. <http://id.nii.ac.jp/1001/00090086/>
- Kobayashi, Yuichiro. “Tahenryokaiseki ni yoru sekai no eigo gakusyusya no danwabunseki [A multivariate discourse analysis of non-native Englishes].” *Jinmonkon2010ronbunshu [Jinmonkon collections of academic papers 2010]* 15 (2010): 41–48. <http://id.nii.ac.jp/1001/00090106/>
- Ministry of Education, Culture, Sports, Science and Technology. “Chugakko gakusyushidoyoryo [The guidelines for the course of study of the junior high school].” 2017. https://www.mext.go.jp/content/1413522_002.pdf
- Ministry of Education, Culture, Sports, Science and Technology. “Kotogakko gakusyushidoyoryo [The Guidelines for the course of study of the high school].” 2018. https://www.mext.go.jp/content/1384661_6_1_3.pdf
- Mita, Kaoru and Shimoda Atsuko. “Gakusei no eibun writing ryoku kozo no bunseki sono2: bunpo/kozo/ronri no zyutenteki shido ni yoru writing no syuzyukudobetsuhenka [Analysis of students’ improvement in English writing. part 2: changes in writing proficiency with focused instruction in grammar, organization, and logic].” *Jissen English communication* 51 (2021): 12–46. <http://doi.org/10.34388/1157.00002251>
- Miyasako, Nobuyoshi. “An analysis of high school student compositions in the sentence-based approach.” *Annual Review of English Language Education in Japan*, 11 (2000): 111–120. https://doi.org/10.20581/arele.11.0_111
- Nakagawa, Hiroyo. “Tankidaigakusei wo taisyo to shita paragraph writing no shido zissen no sono koka [The effectiveness of paragraph writing activities for college level writers of English].” *Kenkyuronsyu [Journal of Inquiry and Research]* 104 (2016): 117–127. <http://doi.org/10.18956/00007706>
- Sawaya, Yusuke and Suzuki Tomoki. “Eigo writing ni okeru kessokusei to hyoka no kannkeiseimatomari no aru writing wo mezasita zissentekikenkyu- [The relationship between evaluation and cohesion in English writing: Practical research aims for coherent writing].” *Hokkaido eigo kyoikugakkai kiyo [Journal of English Education in Hokkaido]* 15, 0 (2016): 35–54. https://doi.org/10.24675/helesje.15.0_35
- Tono, Yukio. *Eigo totatatsudo shihyo CEFR-J gaidobukku: CAN-DO risutosakusei katsuyo [The CEFR-J handbook: a resource book for using CAN-DO descriptors for English language teaching]*. Taishukan Shoten, 2013.

An Analysis of Association Measures in Collocation Extraction from a Pedagogical Perspective

KOHEI FUKUDA

1. Introduction

Many researchers point out the pervasiveness of collocations in a text and the importance of learning collocations in second language acquisition. Hill (2000: 53) argues that

Collocation is important because this area of predictability is, as we have seen, enormous. Two, three, four, and even five-word collocations make up a huge percentage of all naturally-occurring text, spoken or written. Estimates vary, but it is possible that up to 70% of everything we say, hear, read, or write is to be found in some form of fixed expression.

A vast number of language texts are composed of collocation; therefore, collocation learning is essential for proficient use of language.

Sinclair (1991) proposed two models of how words occur in a language text: *the open-choice principle* and *the idiom-principle*. The open-choice principle sees language texts as the result of a large number of choices where the only restraint is grammaticalness. This model regards each slot in which an individual word is put as an “open slot.” For instance, in a grammatical structure of a transitive verb followed by its object, as in *cause accidents* and *make a cake*, virtually any word can occur in the first slot and the second slot as long as the phrase is grammatically well-formed. On the other hand, according to the idiom principle, there are many more constraints and limitations in the choice of words in a text. The choice of one word determines, to some

extent, the choice of another word. For example, the transitive verb *cause* is usually followed as its object by something bad or unpleasant. Therefore, “*cause accidents*” sounds natural, while “*cause victory*” sounds unnatural. Nation (2013: 483) argues that collocation is an important learning goal because a large amount of language is based on the idiom principle.

Despite its pervasiveness in language texts and its importance in the acquisition of a second language, collocation seems to be often ignored by language teachers and learners in Japan. According to a survey conducted by Kawamura & Ishii (2013), no more than 1.6% of Japanese university students comprehended the concept of collocation, which suggests that few learners in Japan have paid attention to collocation in studying English when they were in junior high school or high school. This is partly because the current Course of Study in Japan does not clearly set collocational competency as a learning objective.

Unlike individual words, it is difficult to choose collocation as an aim of study because teachers themselves simply do not know exactly what collocations are and how they should teach them. Furthermore, since individual words produce a massive number of word combinations in principle, there can be too many collocations for learners to acquire. In order to solve the problem inherent in collocation learning, it is necessary to identify a set of collocations which should be acquired by Japanese learners of English. Koya (2012) argues that it is essential to make a “basic collocations list,” which can contribute to clarifying collocation to acquire for learners, and popularizing collocation learning in English education in Japan. Furthermore, L2 collocations can be learned both by incidental and intentional learning, but intentional learning results in bigger and faster gains (Szudarski, 2017: 212). Given that most Japanese people learn English in EFL environment, where they are not exposed to enough input to incidentally learn collocational competence, intentional learning of collocations are of even greater importance, and for that purpose, a collocations list is necessary.

What is essential for creating a collocations list is a set of criteria for identifying collocations for pedagogical purposes. Criteria for identifying collocations are generally divided into two types; the frequency-based view and the phraseological view (Henriksen, 2013). The frequency-based view is an attempt to identify collocations on the basis of statistical measures which assess collocability, generally known as association measures (AMs), using large corpora. The phraseological view employs linguistical classification criteria, such as the degree of semantic opacity, collocational structure, and substitutability of word elements. Granger & Paquot (2008) claimed that researchers should utilize both of these two views in a well-balanced manner in identifying collocation. Therefore, when making a collocations list for pedagogical purposes, researchers should extract collocations using AMs first, and then screen these collocations using the phraseological view from an educational perspective. Selecting collocations based on the phraseological view, however, is a subjective process and requires enormous effort. Thus, making good and efficient uses of AMs in extracting collocations should enhance the reliability and efficiency of the process of choosing an appropriate set of collocations.

Few attempts have been made to explore how to employ AMs in extracting collocations from large corpora from pedagogical perspectives. The main objective of this paper is to explore how AMs of collocations can be used to extract collocations for the purpose of creating a “collocations list” for Japanese learners of English.

2. Review of Related Literature

This section introduces some previous studies on collocational competence of L2 learners and gives an explanation as to why L2 language users need to learn collocation, and what factors have an influence on L2 learners’ collocational competence. Furthermore, this section explains widely used AMs of collocations which this study will deal with.

2.1 L2 language users' need for collocation competence

Many researchers have investigated the acquisition of collocations by L2 learners, and pointed out the importance of collocation competence for language production and reception. Collocational proficiency enables L2 users to make use of fixed phrases, and therefore strike interlocuters or readers as native-like (Henriksen, 2013). O'Keefe et al. (2007) argues that the use of fixed expressions alleviates the burden on users' cognitive ability when processing language and allows language users to direct cognitive energy into more creative aspect of language use, such as discourse organization and successful interaction. In sum, collocational competence assists L2 learners in communicating in a more natural and creative manner.

2.2 L2 learners' collocational competence

Many studies point out the influence of L1 on L2 collocational competence. Nesselhauf (2003) investigated the use of verb + noun collocation by German learners of English, and suggests that a learners' L1 has an influence on the use of collocations. She drew from her study a conclusion that an explicit instruction of collocations is necessary to enhance learners' proficiency. Granger (1998) explored French L2 learners' use of intensive adverb + adjective collocations (e.g., *completely different*). She found that advanced learners overused certain collocations which were equivalent to their mother tongue. She argued that although learners' unnatural-sounding production of language is generally associated with their lack of prefabricated expressions, it can also be due to their overdependence on certain expressions. Kurosaki (2010) studied the use of verb + noun collocations by Japanese university students, and showed that L1 has an effect on the collocational proficiency of Japanese learners.

Koya (2005) explores the process of the acquisition of verb + noun collocations by Japanese learners of English. She suggests that (1) learners' general vocabulary knowledge correlates with collocational knowledge; (2) knowledge of receptive collocational knowledge is deeper than productive collocational knowledge; (3) productive collo-

cational knowledge is influenced by vocabulary knowledge, semantic opacity, delexicalized verbs, core meaning of nouns and verbs, collocational structure, and L1 equivalence; (4) receptive collocational knowledge is affected by L1 equivalence, delexicalized verbs and core meaning of verbs. She claims that learners at every level should pay attention to collocation and that educators should teach collocation differently to learners at different proficiency levels.

2.3 Association measures

Many statistical methods of measuring collocational strength have been developed. Ishikawa (2008) introduces, as widely used AMs, raw frequency, Dice coefficient, t-score, mutual information (MI), Log-likelihood (LL), z-score, and MI3. Raw frequency refers to the number of times when a certain collocation occurs in a corpus.

T-score of 2 or higher is usually considered a statistically significant combination of words, or collocation. MI is used as a measure which shows to what extent a word has information about another word. MI score of 3 or higher can be interpreted as evidence that the combination of the two words is collocation (Hunston, 2002).

McEnery et al. (2006) state that the most commonly used statistical test is the chi-square test and another commonly used statistical test is Log-likelihood (LL). The chi-square test (χ^2) makes a comparison between the observed values and the expected values. LL also compares the observed values and the expected values.

LL is generally preferred, compared to chi-square because (1) it does not presuppose the minimum expected frequencies, (2) it does not overestimate rare cases, and (3) it is not influenced by corpus size (Leech et. al, 2001).

Whereas MI score puts too much emphasis on rare words, MI3 pays more attention to frequent words. Thus, collocations extracted using MI3 are more useful for language learners at the beginning and intermediate level, while those extracted using MI are interesting for a lexicographic purpose (McEnery et al, 2006).

Ishikawa (2008) classifies five measures (raw frequency, t-score, LL,

Dice coefficient, MI) into three categories. Raw frequency, t-score, LL are grouped into frequency-based measures, which put emphasis on high frequent collocations. MI is categorized into non-frequency-based measures. As MI puts weight on low-frequency words which mostly occur with a certain word, MI tends to extract low-frequency collocations. Dice coefficient lies between the two groups of AMs. According to Evert (2008), LL is the best measure in terms of mathematical statistics. T-score is not based on mathematical reasoning, but still it can be useful as a heuristic measure for collocation identification. He argued that it is important to explore what kind of collocations are extracted by different measures.

2.4 Research Questions

Some researchers utilize AMs so as to create collocations lists for pedagogical purposes. Ackermann & Chen (2013), in an attempt to make a collocations list for EAP (English for academic purpose), set the criteria for extracting collocations from corpora as follows: (a) raw frequency ≥ 1 per million; (b) raw frequency ≥ 0.2 per million in each sub-corpora; (c) MI score ≥ 3 ; (d) t-score ≥ 4 . Koya (2015) explores how to select basic collocations for Japanese learners of English to acquire. In the study, she took the noun *time* as an example, and created a collocations list of “verb + *time*”. In extracting “verb + *time*” collocations from corpora, she employed t-score, z-score, MI, and Log-likelihood.

As is seen in the study stated above, AMs help create collocations lists for educational purposes. However, which association measures should be best applied to selecting collocations for pedagogical purposes are yet to be explored. The consideration of how effectively each association measure extracts pedagogically useful collocations and the comparison between these measures are an essential process of investigating the usefulness and suitability of these measures. Therefore, the research questions of this paper were formulated as follows:

RQ1: Which association measure can extract collocations of ped-

agogical use more effectively?

RQ2: How should each association measure be combined to obtain collocations depending on different proficiency levels of learners?

3. Method

In order for language policymakers and practitioners to incorporate collocation learning into a classroom, a collocations list is necessary, as is a word list for vocabulary learning. Extracting collocations from corpora is an essential process of creating a collocations list, and AMs should be employed so as to obtain collocations in an efficient way and on an objective scale. Therefore, the main purpose of this paper is to explore how AMs, such as Dice coefficient, Log-likelihood, t-score, z-score, MI, and MI3, should be used to extract collocations from corpora with pedagogical applications in mind.

3.1 Materials and corpora used in the study

In this study, it was hypothesized that collocations which appear in a published study book for collocations are presupposed to be those of high pedagogical value. Although there are several study books for collocations in Japan and the rest of the world, one of the most widely used is *English Collocations in Use -intermediate (second edition)* (McCarthy & O'Dell, 2017). The present study used collocations found in this book in order to explore the validity of association measures.

The study book is organized into 60 two-page units. Collocations are presented in typical contexts, and each unit focuses on a certain topic, such as weather, music, sport, business, money, time, talking about success and failure, and so forth so that you can learn collocations in a meaningful context. The right-hand page provides a series of exercises so you can check that you have understood the collocations you've studied on the left-hand page.

McCarthy & O'Dell (2017) pay attention mainly to two things when selecting collocations which would be most useful for learners to

study. The first thing is that the authors of the book put emphasis on the collocations which many users of English are likely to use in their speech or writing. “So, in the unit on Eating and drinking we include, for example, *have a quick snack* and *processed food* but not *cocoa butter*, which is a very strong collocation, but one which has very limited use for most people” (McCarthy & O’Dell, 2017: 4). Second, the authors carefully selected semantically opaque collocations, which learners of English might have difficulty in decoding, based on the analysis of the Cambridge Learner Corpus.

In this study, information about AMs was obtained by using the British National Corpus (BNC). The present study made use of BNCweb (Sebastian & Evert) because this interface enables users to extract collocations automatically by seven measures of association, including raw frequency, Dice coefficient, Log-likelihood, t-score, z-score, MI, MI3. Therefore, BNCweb is suited for the current study.

3.2 Corpus processing and data analysis

This study focuses on verb + noun collocations. Verb + noun collocations were listed up from McCarthy & O’Dell (2017), and 637 collocations were identified. Verbs were considered to be a node word, and nouns were viewed as its collocate. If verbs appeared in more than nine collocations, they were selected as target verbs. Those collocations in which the target verbs were used were chosen as an object of investigation. As a result, the following eleven verbs were selected: *get*, *do*, *have*, *give*, *take*, *make*, *keep*, *win*, *raise*, *change*, and *cause*. In total, 250 collocations were identified for these verbs from McCarthy & O’Dell (2017), which was called the “target collocations” in this study (see Table 1).

After selecting verbs as node words to investigate in this study, collocates for those node words were extracted from the BNC by employing the six association measures (AMs), Dice coefficient, Log-Likelihood, t-score, z-score, MI3, and MI. The collocation search span was set to +4 (within the four words in the right context), and collocations were extracted and listed as lemmas. As for each node word, the top

Table 1. The list of “verb + noun” target collocations investigated in this study

VERB + NOUN	Number	Example
do + NOUN	35	do activities, do aerobics, do an assignment
make + NOUN	45	make a breakthrough, make an allegation
have + NOUN	40	have a think, have a break, have a conversation
give + NOUN	25	give credit, give the impression, give a laugh
cause + NOUN	11	cause damage, cause concern, cause pain
change + NOUN	11	change doctors, change jobs, change the subject
win + NOUN	11	win respect, win case, win praise
get + NOUN	10	get a job, get a place, get the impression
keep + NOUN	10	keep the pace, keep a record, keep secrets
raise + NOUN	9	raise a question, raise money, raise taxes

100 collocations were extracted from the BNC by using association measures.

To explore the validity of respective AMs for pedagogical purposes, the author investigated the extent to which the top 100 noun collocates from the BNC matched the target collocations selected from McCarthy & O’Dell (2017). Moreover, the author examined the rank of the target collocations by each association measure in question. Different measures returned different results and values, and therefore the direct comparison across the AMs was not possible as they were. However, the rank order of the target collocations by each AM made it possible to compare the different AMs from one another.

4. Results

The results first show how many of the target collocations were covered by each AM. Second, a comparison was made across the AMs, and it is explored how different measures evaluated collocations, and how they were classified in terms of similarities.

4.1 The coverage of the target collocations by the measures and their rank order

Table 2 shows the coverage of the target collocations by the top 100 collocations extracted by using each AM in question. The results show that on average, the percentage of the target collocations extracted using each association measure were as follows: Dice coefficient: 55%, Log-Likelihood: 56%, t-score: 55%, z-score: 53%, MI3: 59%, and MI: 26%. Approximately 90% of “*cause* + NOUN” collocations were covered, which was extremely high compared to the other; however, only 30% of “*do* + NOUN” collocations were covered on average. Overall, the coverage of MI was much lower than the other five measures.

Table 3 shows the rank orders of the target collocations according to the top 100 collocations list extracted from the BNC by employing

Table 2. The coverage of the target collocations

Node	Total	D	LL	T	Z	MI3	MI	D(%)	LL (%)	T (%)	Z (%)	MI3 (%)	MI (%)
make	45	29	31	29	28	32	9	64%	69%	64%	62%	71%	20%
have	43	19	21	20	21	21	10	44%	49%	47%	49%	49%	23%
take	40	21	23	19	22	23	7	53%	58%	48%	55%	58%	18%
Do	35	12	12	12	12	14	6	34%	34%	34%	34%	40%	17%
give	25	13	14	14	13	14	6	52%	56%	56%	52%	56%	24%
cause	11	10	10	10	10	10	8	91%	91%	91%	91%	91%	73%
change	11	5	5	7	4	5	3	45%	45%	64%	36%	45%	27%
Win	11	11	9	9	9	10	7	100%	82%	82%	82%	91%	64%
Get	10	3	2	3	2	3	0	30%	20%	30%	20%	30%	0%
keep	10	7	7	7	7	7	6	70%	70%	70%	70%	70%	60%
raise	9	8	7	8	4	8	4	89%	78%	89%	44%	89%	44%
total/average	250	138	141	138	132	147	66	55%	56%	55%	53%	59%	26%

Notes: D = Dice coefficient, LL = Log-likelihood, T = t-score, Z = z-score, MI = mutual information

Table 3. The rank orders of the target MAKE + NOUN collocations

Collocate	D	LL	T	Z	MI3	MI
decision	2	3	2	3	3	75
mistake	5	2	6	1	2	15
way	6	20	4	46	11	-
point	7	14	7	27	13	-
contribution	8	6	8	5	6	41
effort	9	7	9	11	7	97
money	11	23	10	40	16	-
progress	12	8	16	9	9	56
change	14	29	15	57	23	-
profit	15	12	19	16	14	-
choice	20	18	24	24	21	-
arrangement	24	16	25	20	18	-
note	26	25	28	30	25	-
comment	28	19	29	21	22	-
impact	33	26	38	29	27	-
friend	36	89	35	-	66	-
start	37	32	40	39	34	-
sound	40	54	42	68	48	-
time	41	-	20	-	76	-
reference	43	50	45	67	50	-
speech	45	48	48	60	52	-
film	49	70	54	88	67	-
demand	51	74	55	99	69	-
appointment	56	49	63	56	55	-
assumption	60	56	66	61	59	-
comparison	64	58	73	62	63	-
case	68	-	59	-	-	-

list	74	-	74	-	90	-
loss	82	-	80	-	94	-
adjustment	-	62	-	53	70	88
allegation	-	-	-	-	-	-
breakthrough	-	-	-	-	-	89
commitment	-	99	-	-	99	-
detour	-	84	-	51	-	20
excuse	-	64	-	55	72	93
headline	-	95	-	92	-	-
improvement	-	-	-	-	-	-
modification	-	-	-	-	-	-
observation	-	88	-	-	100	-
photocopy	-	-	-	-	-	-
preparation	-	-	-	-	-	-
recording	-	-	-	-	-	-
redundant	-	-	-	-	-	-
reservation	-	-	-	-	-	-
withdrawal	-	-	-	-	-	-

the AMs. Numbers indicates the rank orders in each measure, and the unmarked cells (shown by “-“) shows that the target collocations were not found in the top 100 lists. It is obvious that MI covers a very small numbers of the target collocations, and thus it seems that MI is not suitable for identifying pedagogically useful collocations.

4.2 A comparison between the AMs

In Table 3, it seems that the five measures except MI produced apparently similar results. Thus, it is necessary to further explore the differences between the five measures. To this end, the target collocations were compared in terms of the coverage across different AMs. Out of 250, 99 target collocations were covered in all the AMs except

MI. Ten collocations were covered by four measures, twenty-five collocations by three measures, fourteen collocations by two, eight collocations by only one measure.

MI3 covered all the target collocations covered by four or three measures (see Table 4 and Table 5). In table 4, no more than three of the target collocations covered by the four measures were extracted by z-score. As is shown in Table 5 and Table 6, Dice coefficient and t-score produced similar results, and Log-Likelihood and z-score assessed collocations in a similar way.

This result suggests that Dice coefficient and t-score can be grouped together in terms of collocation selection behavior, and Log-likelihood and z-score can be classified into another group of association measures. To explore the difference between Dice and t-score versus LL and s-score, it can be useful to examine the collocations which were extracted by Dice and t-score, but not by LL and z-score and which are covered by LL and z-score, but not by Dice and t-score in terms of word level of collocates.

The pedagogical importance should be assessed by the frequency of

Table 4. The target collocations covered by four out of the five measures except MI, and their rank orders

node	collocate	D	LL	T	Z	MI3
make	friend	36	89	35	-	66
change	place	49	74	29	-	42
raise	subject	63	85	48	-	66
raise	capital	63	85	48	-	66
raise	child	66	92	29	-	52
give	performance	68	94	73	-	84
win	praise	71	57	-	59	64
give	talk	76	85	87	-	90
take	pleasure	99	70	-	89	81
give	sigh	-	47	29	52	45

Table 5. The target collocations covered by three out of the five measures except MI, and their rank orders

node	collocate	D	LL	T	Z	MI3
have	child	13	-	11	-	23
do	course	16	-	16	-	22
have	word	30	-	30	-	45
do	research	37	-	36	-	46
get	place	40	-	37	-	64
take	course	41	-	34	-	70
make	time	41	-	20	-	76
have	view	52	-	54	-	62
do	duty	63	-	67	-	50
make	list	74	-	74	-	90
do	hair	77	-	81	-	66
make	loss	82	-	80	-	94
raise	family	86	-	46	-	75
have	game	92	-	94	-	95
take	clothes	97	85	-	-	91
win	case	98	-	37	-	88
do	washing	-	14	-	17	33
do	cooking	-	20	-	26	53
give	go-ahead	-	32	-	7	26
keep	temper	-	54	-	47	78
make	adjustment	-	62	-	53	70
make	excuse	-	64	-	55	72
have	ability	-	77	-	86	84
take	prisoner	-	86	-	96	98
have	option	-	87	-	96	96

Table 6. The target collocations covered by two measures out of the five measures except MI, and their rank order

node	collocate	D	LL	T	Z	MI3
make	case	68	-	59	-	-
keep	word	90	-	62	-	-
give	word	93	-	79	-	-
do	ironing	-	24	-	24	-
have	chat	-	35	-	25	-
have	tendency	-	52	-	58	-
have	sympathy	-	54	-	56	-
do	exam	-	56	-	70	-
make	detour	-	84	-	51	-
make	observation	-	88	-	-	100
take	trip	-	89	-	-	97
make	headline	-	95	-	92	-
take	photo	-	96	-	99	-
make	commitment	-	99	-	-	99

collocates because the more frequent words are, the more important they are for learners. Therefore, it is meaningful to examine the frequency of collocates listed in Table 7 and Table 8. The frequency of collocates is based on the BNC. It is also valuable to investigate the level of collocates on the basis of CEFR-J Wordlist, because the wordlist is created in order to display the levels of words from an educational perspective (Tono, 2013). Referring to the wordlist enables researchers and educators to know objectively how useful individual words are for Japanese learners of English.

As is indicated by Table 7 and Table 8, the average frequency of collocates that make up the target collocations which are extracted by Dice and t-score is much higher than that by LL and z-score. On top of that, most of the collocates by Dice and t-score fall into A1, while more than half of the collocates by LL and z-score are B1 or on a

Table 7. The target collocations covered by Dice and t-score, but not by LL and z-score, their rank orders, and the frequency and CEFR-J level of their collocates.

node	collocate	D	T	freq	CEFR-J
have	child	13	11	69271	A1
do	course	16	16	56036	A1
have	word	30	30	42301	A1
do	research	37	36	25531	A2
get	place	40	37	52469	A1
take	course	41	34	56036	A1
make	time	41	20	180243	A1
have	view	52	54	30686	A2
do	duty	63	67	11648	B1
make	case	68	59	63148	A1
make	list	74	74	13661	A1
do	hair	77	81	14100	A1
make	loss	82	80	15261	B1
raise	family	86	46	41889	A1
keep	word	90	62	42301	A1
have	game	92	94	20601	A1
give	word	93	79	42301	A1
win	case	98	37	63148	A1

NOTES: freq = frequencies of the collocates in the BNC

higher level.

4.3 Summary

It is clear that the coverage of MI was much more restricted than that of the other five measures, and therefore MI did not seem to be suitable for selecting collocations for pedagogical purposes. In terms of coverage, the five collocational measures except MI yielded a simi-

Table 8. The target collocations covered by LL and z-score, but not by Dice and t-score, their rank orders, and the frequency and CEFR-J level of their collocates.

node	collocate	LL	Z	freq	CEFR-J
do	washing	14	17	1504	-
do	cooking	20	26	1540	A2
do	ironing	24	24	178	B1
give	go-ahead	32	7	271	-
have	chat	35	25	944	B1
have	tendency	52	58	3582	B1
have	sympathy	54	56	2304	B1
keep	temper	54	47	1264	B1
do	exam	56	70	1584	A2
make	adjustment	62	53	2109	B2
make	excuse	64	55	2190	A1
have	ability	77	86	10378	A2
make	detour	84	51	238	-
take	prisoner	86	96	4507	B1
have	option	87	96	9138	B1
make	headline	95	92	1378	B1
take	photo	96	99	2011	A1

lar result, but closer scrutiny revealed that they were classified into three groups. Dice coefficient and t-score seem to make one group, which tend to represent relatively frequent collocations, and Log-likelihood and z-score can form another group, which ranks highly the collocations whose collocates are intermediate level words. MI3 lies between these two groups. MI3 succeeded in extracting many of the target collocations which were covered either by Dice coefficient and t-score, or by Log likelihood and z-score.

5. Discussion

This section summarizes the major findings of this study, and discusses the findings in light of previous studies. Moreover, it will address the question of how the AMs should be used for extracting collocations for educational purposes.

The results of the present study show that MI-score covers much less of the target collocations than the other five association measures. The findings clearly indicate that MI-score is not suitable for selecting collocations from a pedagogical perspective. This result corresponds to McEnergy et al. (2006).

Z-score extracted only three of the collocations which were retrieved by the other four AMs except MI, and MI3 extracted all the collocations. In other words, z-score could not extract the collocations which the other measures rated highly. In addition to that, given the fact that z-score extracted only 53% of the target collocations, which was lower than the other four measures except MI, z-score seems to be less suitable for extracting collocations for pedagogical purposes. On the other hand, MI3 succeeded in extracting all the target collocations covered by the other four or three measures, which means that MI3 can reliably evaluate the collocations that other collocational measures rank highly. Therefore, MI3 seems to be the best measure in selecting collocations which are educationally valuable if you try to employ a single measure instead of combining two or more measures.

The investigation into the target collocations which were covered by two or three measures reveals that Dice coefficient and t-score assess collocations in a similar way, and Log-likelihood and z-score produce a similar result. The collocates which comprise the collocations extracted by both Dice coefficient and t-score are much more frequent and fall into more basic levels according to CEFR-J Wordlist than those extracted by Log-likelihood and z-score. These results suggest that Dice coefficient and t-score are suitable for selecting collocations for learners at an elementary level, while Log-likelihood and z-score are appropriate to the needs of learners at an intermediate or advanced level. This result is inconsistent with Ishikawa (2008), who classified

Log-likelihood and t-score into the same group on the basis of the correlation with raw frequency of collocations.

Since there is no consensus on how to judge collocations by using association measures, it is necessary to examine how the AMs can be combined to select collocations for educational purposes. Three suggestions can be made from the results of this study.

Firstly, the combination of Dice coefficient and t-score can be utilized so as to extract collocations for learners at an elementary level. Dice coefficient and t-score put emphasis on collocations whose components are relatively frequent in corpora and therefore on a more basic level in terms of the CEFR-J. Secondly, Log-likelihood and z-score enable researchers and educators to select collocations which are useful for learners at an intermediate or advanced level. These two measures place relatively higher value on collocations whose collocates are relatively infrequent, and many of the collocates of the target collocations which can be extracted only by these two statistics fall into B1 level or higher on the basis of the CEFR-J. Thirdly, MI3 can be said to be a well-balanced association measure of collocation. MI3 covered many of the target collocations extracted by the four measures except MI. Therefore, MI3 can be an efficient measure when attempting to select collocations which are worth learning for a wide range of students.

It is important to consider how to apply these findings to create a collocations list. Since there are potentially a vast number of collocations, it is vital to reduce the number of collocations in a collocations list to manageable numbers by selecting pedagogically relevant collocations depending on different proficiency levels of learners.

Based on the characteristics of AMs found in this study, the author would suggest a method of selecting collocations. The first step is to use Dice, t-score, and MI3, extract the top 100 lists respectively (this number can be changed according to the size of the intended collocations list), and identify collocations found in all of the three lists. These collocations are meant for learners at a basic level. The second step is to extract the lists using LL, z-score, and MI3. Collocations

extracted by these three measures are intended for lower-intermediate learners. The final step is to identify collocations found in the lists extracted by LL and z-score, not by MI3, and these collocations are meant for upper-intermediate learners. This is just one of the potential methods of creating a collocations list for learners at a basic or intermediate level by using association measures. How to use different association measures is open to discussion. We also have to consider the phraseological approach to collocation and the effect of learners' L1 on the acquisition of collocation in order to select pedagogically relevant collocations.

6. Conclusion

This study aimed to explore how association measures can be utilized in extracting collocations from corpora from a pedagogical perspective. It is important to conduct this kind of research because finding methods to extract educationally useful collocations by using statistics is necessary for making a collocations list efficiently and objectively.

This study explores the usefulness and characteristics of six association measures (Dice coefficient, Log-likelihood, t-score, z-score, MI3, and MI) by investigating how many of the collocations which appeared in a learning book for collocation, *English Collocation in Use*, were covered by the top 100 lists extracted from the BNC using each of the six AMs.

The result of this study suggests that MI is not appropriate for selecting collocations for pedagogical purposes, as was expected from previous studies. A significant finding is that the other five measures in question can be classified into three groups. The first group includes Dice coefficient and t-score, which could be suitable for selecting collocations for learners at an elementary level. The second group is composed of Log-likelihood and z-score, which can be useful in extracting collocations for intermediate learners. MI3 lies between these two groups, and this seems to be a well-balanced measure which can be appropriate for choosing collocations for a wide range of learn-

ers.

The present study has some methodological limitations. First, this study regards collocations which appeared in *English Collocations in Use* as target collocations. Future research should consider collocations which are not included in the textbook. Second, the collocation patterns which this study dealt with were limited to “verb + noun” collocations only. Other collocational patterns, such as “adjective + noun,” “adverb + verb,” and “adverb + adjective” should also be explored because different collocational patterns could yield different results.

Although this study provides a clue as to how AMs can be used for selecting collocations for educational purposes, it remains to be seen how this result should be used in creating a collocations list for Japanese learners of English. Further research will be needed to investigate how AM-derived collocations list can be applied to actual classroom practice while considering other pedagogical factors such as cognitive or affective domains of Japanese learners of English.

REFERENCES

- Ackermann, Kirsten, and Yu-Hua Chen. “Developing the Academic Collocation List (Acl) – a Corpus-Driven and Expert-Judged Approach.” *Journal of English for Academic Purposes* 12.4 2013: 235–47. Print.
- Evert, Stefan. “Corpora and Collocations.” *Corpus Linguistics. An International Handbook*. Ed. Anke Lüdeling, Merja Kytö. Berlin: Mouton de Gruyter, 2008. 1212–48. Print.
- Granger, Sylviane. “Prefabricated Patterns in Advanced EFL Writing: Collocations and Formulae.” *Phraseology: Theory, Analysis and Applications*. Ed. Cowie, A.: Oxford University Press, 1998. 145–60. Print.
- Granger, Sylviane, and Paquot Magali. “Disentangling the Phraseological Web.” *Phraseology: An Interdisciplinary Perspectives*. Ed. Granger, S. & Meunier, F: Benjamins, 2008. 27–49. Print.
- Henriksen, Birgit. “Researching L2 Learners’ Collocational Competence and Development - a Progress Report.” *L2 Vocabulary Acquisition, Knowledge and Use*. Ed. C Bardel, B Laufer & C Lindqvist: Erosia, 2013. 29–56. Print.
- Hill, Jimmie. “Revising Priorities: From Grammatical Failure to Collocational Success.” *Teaching Collocation: Further Development in the Lexical Approach*. Ed. Lewis, Michael: Language Teaching Publications, 2000. 47–69. Print.
- Hoffmann, Sebastian, and Evert Stefan. “BNCweb (CQP-Edition).” Web. <http://corpora.lancs.ac.uk/BNCweb/>
- Hunston, Susan. *Corpora in Applied Linguistics*. Cambridge University Press, 2002. Print.
- Ishikawa, Shinichiro. *Eigo corpus to genjo kyouiku [English Corpus and Language Education]*. Taisyukan, 2008. Print.

- Kawamura, Akihiro and Ishii Yasutake. "Communication Nouryoku no tameno Goi Shidou. [Developing students' vocabulary knowledge for better communication skills : with special emphasis on politeness and collocation]" *Seijo University, Social innovation studies* 8.2, 2013: 37–68. Print.
- Koya, Taeko. "The Acquisition of Basic Collocations by Japanese Learners of English." Waseda University, 2006. Print.
- Koya, Taeko. "Eigo Kyouiku to Collocation [English Education and Collocation." *Korekara no Collocation Kenkyuu [Future Research on Collocation]*. Ed. Hori, Masashiro: Hitsuji Shobou, 2012. 23–60. Print.
- Koya, Taeko. "Kihon Collocation list Sakusei no tameno Ichi Kousatu [A study for creating basic collocations list]." *Hosei University Koganei Ronsyu* 11, 2015: 19–32. Print.
- Kurosaki, Shino. "An Analysis of Japanese L2 Learners' Knowledge of "Verb + Noun" Collocations." *Jissen Women's University FLC journal* 5, 2010: 63–77. Print.
- Leech, Geoffrey, Rayson Paul and Wilson Andrew. *Word Frequencies in Written and Spoken English Based on the British National Corpus*. London: Routledge, 2001. Print.
- McCarthy, Michael and O'Dell Felicity. *English Collocations in Use Intermediate [Second Edition]*. Cambridge University Press, 2017. Print.
- McEnery, Anthony, Xiao Richard and Yukio Tono. *Corpus-Based Language Studies: An Advanced Resource Book*. Routledge, 2006. Print.
- Nation, I.S.P. *Learning Vocabulary in Another Language*. Cambridge University Press, 2013. Print.
- Nesselhauf, Nadja. "The Use of Collocations by Advanced Learners of English and Some Implications for Teaching." *Applied Linguistics* 24, 2003: 223–42. Print.
- O'Keeffe, Anne, McCarthy Michael and Carter Ronald. *From Corpus to Classroom: Language Use and Language Teaching*. Cambridge University Press, 2007. Print.
- Sinclair, John. "Corpus, Concordance, Collocation." *Modern Language Journal* 78, 1991. Print.
- Szudarski, Pawe. "Learning and Teaching L2 Collocations: Insights from Research." *TESL Canada Journal* 34, 2017: 205–16. Print.
- Tono, Yukio (ed). *Can-Do list Sakusei Katsuyou Eigo Toutatsudo Shihyou CEFR-J Guide book [The Creation and Utilization of Can-Do lists CEFR-J Guidebook]*. Taishukan, 2013. Print

投 稿 規 定

(1) 投稿は岩崎研究会会員に限る。但し、非会員であっても論文審査委員から推薦のあった場合は特別に認める。(2) 論文の内容は未発表のものに限る。(3) 用語は英語に限り、原則として native check を受けたものとする。(4) 注 (note) は後注とし、章ごとに通し番号を付ける。(5) ギリシャ字、ロシア字以外の特殊文字はできるだけローマ字化してほしい。音声記号は国際音声学協会 (IPA) 所定のものを用いる。(6) 引用文献：書式は MLA Style に従う。(7) 枚数：論文はワープロ原稿で、1行はアルファベットの小文字で 70 字、450 行以内。(8) 原稿はすべて論文審査委員による審査の上採否を決定する。共同執筆論文を別として、論文の掲載は毎号 1 人 1 篇とする。(9) 都合により短縮を求めることがある。印刷上の体裁および論文の掲載年度については編集委員に一任する。(10) 抜刷は 20 部までを無料で、別に本誌 1 部を呈上する。(11) 原稿は随時受付ける。(12) なお、詳細は別に定める。

論文審査委員

赤須 薫	土肥 一夫
岩崎 春雄	馬場 彰
浦田 和幸	斎藤 弘子
増田 秀夫	清水あつ子
宮井 捷二	投野由紀夫
山田 茂	

編集後記 今年、岩崎研究会は還暦を迎えます。創立二十五周年記念号(1987年)の編集後記で小島義郎先生は「岩崎研究会が発足したのは昭和 37 年 9 月で…四半世紀前に竹林滋氏と私のほか横山一郎氏、若林俊輔氏、小川繁司氏の 5 名で岩崎先生のお宅をお借りして始まったこの会が現在会員 100 名を超える大きな会に育ったことを思うと感無量なものがある。」と述べておられる。わずか 5 人の集まりから、我が岩崎研究会が始まったとは驚きです。みなさんと一緒に研究会の還暦をお祝いしたいと思います。

しかし、喜んでばかりとは行きません。一番気がかりなのは Lexicon への応募論文の数が極端に少ないことです。今回は、赤須先生はじめ他の先生方の尽力で何とか論文が集まりました。本誌に掲載される論文は辞書学に関する論文ばかりではありません。理論言語学、応用言語学、英語学、外国語教育などの分野に対しても広く戸門を開くものであります。これからも、多くの会員、特に若い研究者がふるって論文を発表されることを期待します。

(2022 年 5 月 30 日 S. M.)

