〔論 文〕

Effective Methods of Teaching "Verb + Noun" Collocations: Japanese EFL learner subcorpus-based analyses

Takako KOBAYASHI

Abstract

In the present study I investigate the use of lexical collocations in Japanese EFL learners' written English, with a special focus on the verb "take." I compare differences in the use of "take" in academic essays written by Japanese university students and by American university students, using the ICLE Japanese subcorpus as a learner corpus and LOCNESS as a reference corpus. I examined lexical collocations of "take" in terms of frequency, variety, overuse, and underuse. The results revealed that although the total frequency of "take" with noun combinations was higher in the Japanese EFL learner data, learners tended to overuse fixed expressions and did not use native-like collocations. The overall results suggest the need to expand collocational knowledge on the part of Japanese EFL learners, which may prove to be beneficial in producing more natural sentences. To achieve the above goal, I propose three suggestions in this paper. The first is that the process by which Japanese EFL students learn collocations needs to be clarified. The second is that collocational data retrieved from corpora should be put to use in classroom English texts. Lastly, there should be a greater awareness of the importance of teaching collocations to Japanese learners of English.

Key words: collocations, high-frequency verbs, ICLE Japanese subcorpus, LOCNESS, use of natural English in writing

1. Introduction

The purpose of this paper is to explore effective methods for teaching collocations to Japanese EFL learners. This issue is particularly relevant given the heightened interest in recent years in collocations, which are forms in which combinations of individual lexical items frequently co-occur. Collocations have important implications in that they comprise co-occurrence information that exists between lexical items.

Although many scholars have noted the importance of collocations, in this paper I first explore what constitutes an English "collocation," an issue that has yet to be resolved. I then examine what the "basic" collocations are for Japanese EFL learners. Next, I attempt to shed light on methods that might be effective for teaching collocations to these learners. Furthermore, through analyses based on a comparison of an actual Japanese EFL learner subcorpus and a corpus of writings by native English speakers, I examine differences seen in the use of collocations between both groups.

In recent years, corpus-based research has advanced considerably and become an even

more pervasive tool. At the same time, a growing number of researchers have been calling attention to the importance of collocations. As a result, interest in collocation learning has grown from an educational standpoint. It has become increasingly common to see analyses utilizing learner corpora, comparisons of learner corpora with those of native speakers, and examinations of potential teaching methods with foundations in such findings.

Meanwhile, with regard to collocation instruction, there are many scholars that invoke the traditional or conventional properties of collocations. There is still a strong tendency to view collocations as the hallmark of the proficiency of native speakers of English. Put differently, while some recognize the importance of and advocate for collocation instruction, others still take the position that the production of collocations and the determination of their appropriateness are the exclusive purview of native speakers. In this paper, I take into account these conflicting views on collocations, yet the purpose of my research lies in the conviction that researching the properties of collocations of English as the "Global Language" will contribute to the future of English education in Japan.

For Japanese EFL learners, in order to be able to compose native-like English, it is necessary to learn the correct ways in which nouns and verbs connect. It is very common for something that would be normally be expressed by a "verbal phrase" in Japanese to be expressed by a "noun phrase" (a verb and a noun) in English. Put differently, Japanese is more verb-centric, while English is more noun-centric. For this reason, for Japanese learners of English to develop the ability to write natural English, they must inevitably master "verb + noun" collocations. This study is dedicated to the examination of such "verb + noun" collocations.

2. Literature Review

Collocation generally refers to the phenomenon of particular words commonly juxtaposed. In Japanese, it is often translated as *rengo-kankei*. Stubbs (2002) states that "Collocation refers to individual word-forms, which are directly observable in texts" (p. 88).

In recent years there has been a substantial increase in the volume of research on collocation, a topic which is not without issues. Some of these issues include the ambiguity of the definition of "collocation," the lack of clear distinctions between collocations and idioms or free combinations, and the absence of an agreed-upon vocabulary for discussing collocations.

In English education settings in Japan as well, a variety of terms are used to refer to collocations, such as rengo (a technical rendering of "collocation"), jukugo ("set phrase"), kanyōku ("idiom"), and kōgo-hyōgen ("colloquialism"). Despite the many assertions that the study of collocations is important for improving English ability among EFL learners, the fact remains that Japanese students of English have not always received systematic collocation instruction.

Sinclair (1991) remarks on the difficulty of collocation and idiom use among non-native English speakers as follows:

At present, many learners avoid the common words as much as possible, and especially where they make up the idiomatic phrases. Instead of using them, they rely on larger, rarer, and clumsier words which make their language sound stilted and awkward. (p. 79)

Lewis (2000) defines collocation as "the way in which words co-occur in natural text in statistically significant ways" (p. 132).

The importance of collocation instruction is often discussed in the context of arguments involving memory, language fluency, appropriateness of use, lexical models, and the effectiveness of instruction. According to the notion put forward by Sinclair (1991) that most words are used in a collocated form, depending upon how text is composed, there is at work an open-choice principle and idiom principle. Language users therefore have at their disposal a large number of semi-preconstructed phrases. For this reason: a) there are limitations to co-occurring words, b) frequent words tend to be delexicalized, and c) such usage is guided predominantly by the idiom principle.

Benson et al. (2009) go further in that they not only point out the necessity of learning collocations, they also make the distinction between grammatical collocations and lexical collocations.

Students must learn how words combine or 'collocate' with each other. In any language, certain words regularly combine with certain other words or grammatical constructions. These recurrent, semi-fixed combinations, or collocations, can be divided into two groups: grammatical collocations and lexical collocations. (p. xiii)

A grammatical collocation is defined by Benson et al. (2009) as follows:

A grammatical collocation is a phrase consisting of a dominant word (noun, adjective, verb) and a preposition or grammatical structure such as an infinitive or clause. (p. xix)

For example, "decide on" in "decide on a boat" and "adapt to" in "adapt to new conditions" are combinations of content and function words, and are intricately linked to grammatical and syntactic conventions.

On the other hand, Benson et al. (2009) define lexical collocations as follows:

Lexical collocations, in contrast to grammatical collocations, normally do not contain prepositions, infinitives, or clauses. Typical lexical collocations consist of nouns, adjectives, verbs, and adverbs. (p. xxxi)

According to Benson et al. (2009), "lexical collocations" consist of "verb + noun," "adjective + noun," or "noun + verb" combinations, each element of which is equal to the other. An example given is "warmest regards" as in "I send warmest regards." They also assert that most lexical collocations comprise "verb + noun" combinations.

Lewis (2000) recognizes the importance of the collocation as a concept in the classroom setting, but also notes the importance of learning combinations in their collocated—or "chunk"—form, rather than as individual words.

In most classes learners will already know many individual words, so in these circumstances, they may need to learn about 'putting them together' in standard collocations, but this is part of the necessary artificiality of language teaching. It would unquestionably be better if learners had acquired the words together as a single chunk—a single choice—in the first place. (p. 132)

Following Benson et al. (2009), in the present paper I make use of the distinction between grammatical and lexical collocations. Here I examine combinations consisting of a verb followed by its object, specifically lexical collocations of the verb "take" followed by a noun, which should occur with great frequency in learner corpora.

3. Method

In this paper, to sample verbs frequently used as collocates in the pattern of "verb + noun," I referred to the Longman Grammar of Spoken and Written English (LGSW), which was compiled from a corpus of around 40 million words. According to the LGSW (1999) itself: "Three verbs are particularly productive in combining with a following noun phrase to form relatively idiomatic expressions: have, take, and make" (p. 1026). These verbs can take on a great many meanings. There are also tremendous possibilities in terms of the possible objects with which these verbs can collocate. Because by themselves these are verbs that lack discreet meanings, I predicted that there would be observable discrepancies in their use in collocations among non-native English speakers and native English speakers. Based on this premise, in this paper I focused on examples of "take + noun" as collocations likely to occur with a high frequency often misused by among Japanese learners of English, and compared and analyzed the usage of "take + noun" collocations by this group against native speakers of English

Specifically, I sampled examples of "take + noun" as collocations from a Japanese EFL learner corpus, as I expected this pattern to occur with a particularly high frequency, and analyzed those examples. I investigated the kinds of nouns that typically co-occurred with "take," and considered the characteristics of those combinations as well as examples of their misuse. I also conducted a similar analysis of data from a corpus of native English speakers. In the context of these analyses, I address issues that Japanese learners of English face when tasked with producing "native-like" written English compositions.

Furthermore, in the course of examining the view that proper collocation use is unique to native speakers, I used native English speaker corpora that enabled myself to compile subjective data on the use of collocations by native English speakers, which I used for quantitative analysis.

4. Corpora Used

In this study, the International Corpus of Learner English (ICLE), which is compiled from data that draws upon written sources by learners of English from a wide variety of mother tongue backgrounds will be used. The ICLE is a project that was begun in 1990 by the Louvain Centre for English Corpus Linguistics (CECL) at Belgium's Université catholique de Louvain. At present, there are 19 partner countries.

The ICLE is compiled from argumentative essays totaling at least two million words, each essay of which is at least 500 words in length and composed by an advanced learner of English (juniors and seniors at university) from one of these partner countries. The purpose of this corpus is to serve as a resource for comparative studies of the grammar, vocabulary, and discourse of English as an intermediate language used by learners of different mother tongues. An "interlanguage" is a term for form of a language unique to language learners that is structurally somewhere between the native tongue and the target language. It is a transitional language form that occurs when people learn new languages. The team that is responsible for the compilation of the ICLE Japanese subcorpus belongs to Showa Women's University. The data which they finished collecting was incorporated into the Japanese subcorpus of Version 2 of the ICLE, which was published in July of 2009.

This subcorpus by Japanese learners of English totals 198,241 words, and consists of argumentative essays of at least 500 words by 366 people. In the following, I refer to this data as the "Japanese ICLE data." These essays are written on a wide variety of subjects. It should also be noted that there were no time restraints or restrictions on dictionary use.

The LOCNESS (Louvain Corpus of Native English Essays) was compiled based on the same criteria as ICLE, and comprises argumentative essays by native English speakers totaling 300,000 words. Put together specifically for the purposes of comparative research, it consists of essays deemed suitable for comparative analyses. In the following, I will use the subcorpus of essays totaling 149,574 words written by American university students, which I refer to hereafter as the "LOCNESS data." The average length of the essays upon which this subcorpus of native English speakers is built is 850 words.

Table 1
The total number of essays and total number of words in the corpus used

Subcorpus type	Total essays	Total words
Japanese ICLE data	366	198,241
LOCNESS data	232	149,574

5. Frequency of "take + noun" Collocations in Learner and Native Speaker Corpora

In analyzing the corpora, I used the analytical tools in the WordSmith software to search these two datasets for usages of *take*, *took*, *taken*, and *taking*. I individually examined each and every example of collocations consisting of "*take* + noun" in the concordance lines. Furthermore, using the "collocates" analytical tool in the WordSmith software, I identified the number of noun variations co-occurring in collocations with *take* as well as the frequency of such collocations in the sets of learner data and native speaker data (Table 2).

Table 2 Overall frequency of *take* collocations and frequency of co-occurring nouns

	Japanese ICLE data	LOCNESS data	
take/took/taken/taking	388	297	
"take + noun" collocations	233 (Excluding 14 examples of misuse)	142	
Noun variations	42	55	

As shown in Table 2, the respective occurrences of take, took, taken, and taking collocations in the Japanese ICLE data and LOCNESS data are 388 and 297, respectively. Of these, I examined all concordance line examples and identified examples of nouns co-occurring with take. I chose this analytical method based on the belief that when automatically sampling collocation data, the most important factor should be the frequency of occurrence of word combinations.

Of those instances, there were 233 examples in the Japanese ICLE data and 142 examples in the LOCNESS data in which co-occurrences with *take* were deemed to impart specific meanings and to appear with greater frequency than they would by chance, or in other words, collocation relationships existed between the verb *take* and these nouns.

Table 3 shows these as a ratio of occurrence per 100,000 instances. A comparison using a chisquare test of the frequency of "take + noun" collocation occurrences showed a significant
difference ($\chi^2[1] = 4.04$, p = .044). We can infer from the test that among usage examples of take, took, taken, and taking in both corpora, examples of "take + noun" occur more frequently
among Japanese learners than native speakers, which led to the conclusion that Japanese
learners use "take + noun" collocations more frequently than native speakers of English.

Table 3 Frequency (per 100,000 words) of "take + noun" co-occurrences

	Japanese ICLE data	LOCNESS data
"take + noun" collocations	117.53	94.94
Noun variations	21.19	36.77

Table 4
Examples of frequencies (per 100,000 words) of co-occurring nouns

Japanese ICLE data	Frequency	LOCNESS data	Frequency
land(s) [37]	18.66	advantage [11]	7.35
time [13]	6.56	time [9]	6.02
exam/examination [13]	6.56	consideration [6]	4.01
place [11]	5.55	account [3]	2.01
picture(s) [8]	4.04	class(es) [3]	2.01
part [7]	3.53	life [3]	2.01
lesson(s) [5]	2.53	stand [3]	2.01

Note: The numbers in brackets represent the overall frequency of individual nouns. The nouns listed are those that occur at least five times in the 233 examples in the Japanese ICLE data or at least three times in the 142 examples in the LOCNESS data.

When I used learner data to compare the types of nouns frequently co-occurring with take, I found that for the 117.53 instances of "take + noun" collocations per 100,000 words, there were 21.19 variations of co-occurring nouns. Of those, seven nouns co-occurred five or more times with take. In terms of frequency of occurrence per 100,000 words, they were "land(s)" 18.66, "time" 6.56, "exam/examination" 6.56, "place" 5.55, "picture(s)" 4.04, "part" 3.53, and "lesson(s)" 2.53. This revealed an extremely topic-specific, limited set of co-occurring nouns. The use of "land(s)," which was the most frequently co-occurring noun with take, is likely closely related to the essay topic used in the compilation of the Japanese ICLE data. Because there were numerous examples of fairly specific statements to the effect that "Government should not have the right to take land(s) away without permission," one can safely assume that the dataset includes numerous responses to essay topics of a political nature. The Japanese ICLE data consists of compositions written in response to a diverse array of topics, so it is impossible to deny the existence of bias in depending on composition topics. When one analyzes a corpus with a particular focus on collocations, the findings of that analysis is very likely to depend largely on composition topics. Nonetheless, despite this potential problem, for the purpose of comparison and analysis with a corpus of output by native speakers of English, I have chosen to use the Japanese ICLE data, as I deemed it useful for observing trends in collocation usage commonly seen among Japanese EFL learners.

Next, based on the data represented in Table 3 and Table 4, I compared an analyzed types of nouns that frequently co-occur with take in both the Japanese ICLE data and LOCNESS data. Among the 233 examples in the Japanese ICLE data deemed to be examples of collocations of the "take + noun" variety, there were 14 instances of mistaken usage. Specifically, they were: "they can take a new style of education," "almost everyone takes a high school education," "we take a cold," "I want to take a degree," "it is natural for me to take an artificial heart," "we do not want to take danger," "my friends and I are taking different departments," "we can also give and take e-mail," "if I have children, I will make

my children take English early," "we have a lot of good examples for taking English as the official language," "Now it is said that companies take not educational background," "if children go there, they can take such a wonderful education," "if we take this way of thought," and "we have not to take the problem as easy." Most of these examples of take reflect the imposition upon English of various senses of the Japanese verb toru or ukeru, but they are incorrect in the sense that a native English speaker would choose different verbs in each case. As a consequence of this phenomenon, it is highly likely in the "interlanguage" of learners of English, in which take has numerous and ambiguous meanings, there is confusion occurring with similar collocations using different lexical items. This suggests that the native tongues of learners have a strong influence on the collocations they use. As Granger (1998) states, "... we have established that learners are using collocations, but that they underuse native-like collocations and use atypical word-combinations" (p. 152).

Next, I wish to address the findings of the analysis of the LOCNESS data. As stated above, there were 142 examples of what were deemed to be "take + noun" collocations in this corpus of native English speakers. Because there was an especially strong bias among the nouns co-occurring with the verb take in the Japanese ICLE data, in Table 4 I have listed only those nouns which co-occur at least five times. Nonetheless, although the overall frequency of "take + noun" collocations in the LOCNESS data was low, the range of cooccurring nouns was very broad. As a consequence, there were few examples of nouns with at least five or more examples, so in Table 4 I chose to list nouns with at just three or more examples. What is most interesting here is that in terms of overall frequency of "take + noun" collocation use, the LOCNESS data showed a significantly lower frequency than the Japanese ICLE data (p = .044), but despite this, in terms of variations of co-occurring nouns, the trend is reversed, with a frequency of 21.19 for the Japanese ICLE data and 36.77 for the LOCNESS data. A comparison using a chi-square test of the variations in nouns showed a significant difference (χ^2 [1] = 6.88, p = .009). Put differently, with regard to nouns co-occurring with take, this shows that native speakers use a more diverse range of collocations than Japanese learners. This is evidence that, as stated above, Japanese learners tend to rely more on topic-related collocations, and tend not to use the kinds of ordinary, general collocations that native speakers use in essays.

The surveys I have employed in this study have shown that in the ICLE data "take + noun" collocations are used with significantly greater frequency than among data compiled from native English speakers. Nevertheless, it became clear from the concordance lines that there exists a large discrepancy between the two datasets in terms of collocation use and density. Specifically, Japanese EFL learners use "take + noun" collocations at a rate that is far higher than native English speakers, and the differences in usage from native speakers lead to unnatural collocations, but at the same time, it became evident that native English speakers use collocations that employ a far richer vocabulary.

6. Pedagogical Implications

Based on the findings of this study, I wish to offer the following three proposals regarding the future of collocation instruction to Japanese EFL learners:

1) The process by which Japanese EFL students learn collocations should be clarified

In this study I focused on "take + noun" collocations, but it is nevertheless reasonable to conclude that more careful studies of the overall acquisition process of collocations by Japanese learners of English would lead to more effective methods of collocation instruction.

2) Educational materials utilizing corpora should be developed

Collocations and idioms are addressed in existing lexicons, but they seem to be selected based on the experiences of the author or the editor. Hence, by applying the findings of empirical studies such as this in which learner and native speaker corpora are employed, it is this author's wish that educational materials be developed that aim to increase the frequency of use of collocations, offer methods of acquisition that are tailored to suit difficulties in acquisition, the right order in which the language should be learned, and the vocabulary levels on the part of the learner, resulting in an effective acquisition of that language.

3) Educators should have a better understanding of the significance of collocation instruction

As noted in the literature review above, many researchers have stressed the importance of collocation instruction from a variety of perspectives, yet in actual educational settings, the attention given to collocation instruction is not nearly adequate. It is urgent that educators themselves develop a better understanding of the importance of collocations in English education and develop concrete methods for Japanese learners to study and acquire collocations.

7. Conclusion

In the present study, I have investigated how Japanese learners of English use collocations taking the form of "take + noun." This entailed studying the overuse of specific nouns as well as the misuse of other nouns in forming these collocations. In addition, in the course of comparing uses by Japanese EFL learners against data compiled from uses by native English speakers, I found that Japanese learners use "take + noun" collocations with a statistically significant higher frequency than native speakers, yet in terms of the diversity of the nouns used, the data showed different trends between the two groups. More specifically, Japanese EFL learners tend to repeatedly use a set of nouns that is more

limited than what native speakers use. The variety of nouns used in "take + noun" collocations by native speakers of English was significantly higher statistically. In the course of studying uses of "take + noun" collocations by Japanese EFL learners, numerous misuses were discovered that could be traced to the direct misapplication of words in English with similar meanings in Japanese. With a word like take in particular, its ambiguity as a fundamental lexical unit—and as this study has shown its potential to carry many meanings—makes it especially important for learners to have a basic knowledge of the nouns that commonly co-occur with it in order to master more natural English expressions.

Hunston (2002) remarked as follows on the notion of a lexical syllabus as proposed by Sinclair and Renouf (1988).

At its most simple, the argument is that it makes sense to teach the most frequent words in a language first. Sinclair and Renouf argue that 'the main focus of study should be on (a) the commonest word forms in the language; (b) the central patterns of usage; (c) the combinations which they usually form' (p. 148). Their point is that the most frequent words have a variety of uses, so that learners acquire a flexibility of language fairly easily. (p. 189)

Following Hunston's (2002) line of reasoning, in this paper I offered three proposals for directions that instruction should take regarding frequently-used "verb + noun" collocations.

In addition to these proposals, for learners to develop the ability to appropriately use collocations in the target language, they must not simply engage in the laborious task of rote memorization of collocations, they must also explicitly learn the meaning of target collocations in their proper contexts. For Japanese learners to improve their knowledge of English collocations, research should be encouraged over a range of academic areas, including corpus research. The present study was limited in focus to analyses of collocations consisting of *take* and a noun, but in the future I wish to continue research into other lexical collocations that might prove useful in language acquisition and examine this theme in more depth.

References

- Altenberg, B., & Tapper, M. (1998). The use of adverbial connectors in advanced Swedish learners' written English. S. Granger (Ed.), *Learner English on Computer* (pp. 80-93). London: Longman.
- Benson, M., E. Benson, & R. Ilson (2009). *The BBI Dictionary of English Word Combinations* (3rd ed.). Amsterdam and Philadelphia: John Benjamins Publishing Company.
- Biber, D., S. Conrad, & R. Reppen (1998). Corpus Linguistics: Investigating Language Structure and Use. Cambridge: Cambridge University Press.
- Biber, D., S. Johansson, G. Leech, S. Conrad, & E. Finegan (1999). Longman Grammar of Spoken and Written English. Harlow: Pearson Education.
- Cowie, A. P. (Ed.), (1998). Phraseology: Theory, Analysis, and Applications. Oxford: Clarendon Press.

- De Cock, S., S. Granger, G. Leech, & T. McEnery (1998). An automated approach to the phrasicon of EFL learners. S. Granger (Ed.), *Learner English on Computer* (pp. 67-79). London: Longman.
- Granger, S. (1998). Prefabricated patterns in advanced EFL writing: Collocations and formulae. Cowie, A. P. (Ed.), *Phraseology: Theory, Analysis, and Applications* (pp. 145-160). Oxford: Clarendon Press.
- Hatch, E. & C. Brown (1995). Vocabulary, Semantics, and Language Education. Cambridge University Press.
- Hunston, S. (2002). Corpora in Applied Linguistics. Cambridge: Cambridge University Press.
- Kobayashi, T. (2012). The Pedagogical Implications of Multiword Units in ICLEv2. (pp. 1-12). *Gakuen*, No. 862.
- Lewis, M. (Ed.), (2000). Teaching Collocation: Further Developments in the Lexical Approach. Hove: Language Teaching Publications.
- Moon, R. (1997). Vocabulary connections: Multi-word items in English. N. Schmitt, & McCarthy, M. (Eds.), Vocabulary: Description, Acquisition and Pedagogy. Cambridge: Cambridge University Press.
- Moon, R. (1998). Fixed Expressions and Idioms in English: A Corpus-based Approach. Oxford: Clarendon Press.
- Nesselhauf, N. (2005). *Collocations in a Learner Corpus*. Amsterdam and Philadelphia: John Benjamins Publishing Company.
- Sinclair, J. (1991). Corpus, Concordance, Collocation. Oxford: Oxford University Press.
- Sinclair, J. (Ed.), (1987). Collins COBUILD English Language Dictionary (1st ed.). London: William Collins Sons & Co. Ltd.
- Sinclair, J. & Renouf, A. (1988). A lexical syllabus for language learning. R. Carter and M. McCarthy (Eds.), *Vocabulary and Language Teaching*. (pp. 140-158). London: Longman.
- Stubbs, M. (2002). Words and Phrases: Corpus Studies of Lexical Semantics. Oxford: Blackwell Publishing.
- Sugiura, M. (2002). Collocational knowledge of L2 learners of English: A case study of Japanese learners. T. Saito, J. Nakamura, & S. Yamazaki (Eds.), *English Corpus Linguistics in Japan*. (pp. 303-323). Amsterdam: Rodopi.
- Wray, A. (2002). Formulaic Language and the Lexicon. Cambridge: Cambridge University Press.

(小林 多佳子 現代教養学科)