The Uses of *Get* in Japanese Learner and Native Speaker Writing: A Corpus-based Analysis

Yoko SUZUKI

1. Introduction

Corpus-based studies have revealed the existence of highly conventional collocations and demonstrated that collocations often involve basic verbs frequently used in discourse, so-called "high-frequency verbs" (Altenberg & Granger, 2001; Biber, Conrad, & Cortés, 2004; Groom, 2005). Although these verbs are introduced to EFL learners at an early stage and thus are familiar words to them, learners often have problems with these verbs. For instance, Nesselhauf (2004) analyzes EFL learners' use of the verbs *make*, *have*, *take*, and *give*, and Altenberg and Granger (2001) focus on the collocations with the verb *make*. By using learner corpora, these studies compare native speaker writing with non-native speaker writing and point out that EFL learners feel safe with some uses of a high-frequency verb, while avoiding other uses of the same verb.

This paper compares differences in essays written by Japanese learners of English (henceforth JLs) and native speakers of English (NSs) and examines JLs' use of the verb *get*. The research questions in this paper are as follows:

RQ 1: Do JLs tend to over- or underuse the verb *get*?
RQ 2: What types of use differentiate JLs from NSs?

By examining these research questions, this study identifies the characteristics of JLs' use of the verb *get* and explores the reasons why they deviate from NSs' norms.

2. Background

2.1 Characteristics of High-frequency Verbs

Every language has basic verbs that are frequently used in discourse and found in any corpus-based list of high-frequency verbs. According to *Longman Grammar of Spoken and Written English*, the most common English verbs are *say*, *get*, *go*, *know*, *think*, *see*, *make*, *come*, *take*, *want*, *give*, and *mean* (Biber et al., 1999). High-frequency verbs have unique characteristics distinct from other general verbs. For instance, Viberg (1996) characterizes high-frequency verbs from a cross-linguistic perspective: (i) they express basic meanings and tend to dominate different semantic fields; (ii) they have high-frequency equivalents in most languages; and (iii) they are highly abstract and polysemous. Another characteristic of high-frequency verbs is that they are often used in light verb or delexical constructions (e.g., *make noise*, *take a bath*). A light verb is "a verb with little or no semantic content of its own which combines with a (usually indefinite) direct object noun or NP which itself expresses a verbal meaning" (Trask, 1993, p. 160). Although the meaning of the verb is transparent, there are strict restrictions on the range of nouns which can combine with specific light verbs (e.g., *make noise*, but not **make a bath*).

Previous studies have demonstrated that these characteristics of high-frequency verbs pose problems even for advanced EFL learners, although the verbs function as core words in English and thus are usually introduced to learners at an early stage (Altenberg & Granger, 2001; Lennon, 1996).

2.2 Collocations

As Stubbs (2002) claims, "our knowledge of a language is not only a knowledge of individual words, but of their predictable combinations, and of the cultural knowledge which these combinations often encapsulate." Such combinations are often called collocations[1]. Collocations are defined as combinations of words, selected conventionally rather than linguistically (Sinclair, 1991). A number of corpus-based studies have revealed the existence of highly conventional collocations and demonstrated that collocation is an aspect of language problematic for EFL learners

(Biber, Conrad, & Cortés, 2004; Groom, 2005). Competent use of collocations has been considered as an important part of fluent language use (Nattinger & DeCarrico, 1992; Pawley & Syder, 1983). For example, it is often pointed out that NSs or advanced learners tend to use a variety of collocations, since retrieving readymade combinations of words requires less mental effort than composing an utterance word for word (Wray, 2002). As Warren (2005) suggests, some collocations are memorized not only because they are frequent but also because they are associated with a certain salient type of situation or phenomenon. In the case of the transitive verb *drop*, for instance, collocations would include combinations such as *drop a pen, a glass, a key,* or *a piece of amber*, but **drop love* or **drop sunshine* would not be recognized as correct.

Sinclair (1991) posits two distinct interpretative principles for explaining the way in which meanings arise from language texts: *the open-choice principle* and *the idiom principle*. The open-choice principle views the text as the result of a series of complex choices. In this process, syntax specifies the slots into which memorized items—normally single words—can be inserted (Warren, 2005). The idiom principle views the text as constructed from a number of collocations or semi-constructed chunks which are commonly used in discourse. Jackendoff (1997) points out that there are a vast number of memorized expressions, which can hardly be a marginal part of our language.

2.3 High-frequency Verbs in EFL

The literature on the use of high-frequency verbs starts with two seemingly contradictory observations (Altenberg & Granger, 2001: 174). One observation is that EFL learners tend to overuse high-frequency verbs. On the other hand, analyzing texts written by Norwegian learners of English, Hasselgren (1994) finds that even advanced learners heavily rely on core words such as *give*, *get*, *take*, *show*, *have*, *know*, *keep*, *tell*, and *make*. Such familiar words or phrases are called "lexical teddy bears" because learners feel safe with these expressions. She attributes this tendency of overuse to the fact that "core words—learnt early, widely usable, and above all safe (because they do not show up as errors) are hugely overused, even among learners sufficiently advanced to have been weaned off them" (Hasselgren, 1994, p. 250). Källkvist

(1999) and Granger (1996b) have also reported similar overuse tendency for Swedish learners and French-speaking learners. On the contrary, however, Sinclair (1991) makes a distinct observation that EFL learners tend to avoid using high-frequency verbs. Altenberg and Granger (2001) account for this avoidance in terms of the learners' awareness of the difficulty of these verbs. It is because the choice of basic verbs, the choice of *take* in *take a step* (rather than *make*, for instance), is mostly arbitrary, that is, semantically unmotivated (Allerton, 1984) that learners sometimes hesitate in using these verbs, although they are familiar with them.

Altenberg and Granger (2001) provide a plausible explanation for this contradiction. Comparing the uses of the verb *make* in texts written by Swedish learners, French-speaking learners, and NSs, they report that EFL learners feel safe with some uses of *make* (i.e., *make* "to produce something" and causative *make*), whereas they avoid other uses of *make* (i.e., *make* "to earn money" and delexical *make*). This finding confirms the two competing observations discussed above, indicating the complexity of learners' use of the verb *make*. Analyzing essays written by JLs, Mochizuki (2007) also has similar findings. Although JLs overuse the verb *make* and especially idiomatic *make* (e.g., *make it*, *make the most of*), they underuse certain structures such as creative *make*, light verb *make*, phrasal/PP *make* (e.g., *make up*).

This study adopts the same approach as Altenberg and Granger (2001) take, that is, Contrastive Interlanguage Analysis (CIA) (Granger, 1996, 1998), which compares native speaker data with non-native speaker data. This method has highlighted a number of features that characterize learner interlanguages, and corpus-based studies, especially learner corpora studies, have played a critical role. As Borin and Prütz (2004) claim, the learner corpus has become "one of the most important resources for studying interlanguage," and it enables us to analyze the interlanguage of EFL learners from not only a qualitative but also a quantitative perspective. Taking this approach, the present study focuses on another high-frequency verb *get*. *Get* is lower in frequency in native written English, and thus it might not be a major representative of high-frequency verbs. From EFL learners' perspective, however, *get* is likely to function as a lexical teddy bear, since it is introduced early and every learner should

be familiar with this verb. In fact, Ishikawa (2013) reports that the verb *get* is one of the two high-frequency verbs in the list of the top 30 words JLs tend to overuse. Comparing the use of the verb *get* between JLs and NSs, this study aims to identify JLs' characteristics and explore the reasons for the differences.

3. Data and Method

3.1 Data: Participants

I extracted JL and NS data from the International Corpus Network of Asian Learners of English (the ICNALE Version 2.1; Ishikawa, 2013). The ICNALE is a collection of 1.3 million words of essays written by 2,600 college students in 10 Asian countries and areas and by 200 native speakers of English. It is one of the largest learner corpora publicly available and the only learner corpus focusing on various Asian learners. One of the benefits of using the ICNALE is that writing conditions such as topic, time, length, and use of references are strictly controlled, which makes the data as homogeneous as possible. Detailed writing conditions are described in the Appendix. Controlling writing conditions is effective for comparing different writer groups. However, it should be noted that the variations of words in the data tend to depend on the topics of essays. In the case of the ICNALE, learners write essays on two topics: (Topic A) college students having a part-time job and (Topic B) smoking in restaurants. Therefore, some words or phrases related to smoking or part-time jobs such as *smoke*, *college*, *student(s)*, *job*, *quit* are frequently found in the data, while other words unrelated to the topics are not likely to appear.

	JLs	NSs
Number of words	179,042	90,613
Number of essays	800	400

Table 1: *The Number of Words and Essays in the Data*

Table 1 gives the number of words and essays I used for the present study. The JL data contains almost 180,000 words written by 400 JLs. The NS data contains 90,000 words written by 200

NSs. JLs are grouped into four CEFR (Common European Framework of Reference) levels according to their TOEIC or TOEFL test scores. Although the CEFR originally classifies foreign language proficiency into six levels, A1 (Breakthrough), A2 (Waystage), B1 (Threshold), B2 (Vantage), C1 (Effective Operational Proficiency), and C2 (Mastery), Ishikawa (2013) deletes the A1 level, merges B2, C1, and C2 into B2+ and subdivides B1 into B1_1 and B1_2 in order to describe Asian learners' variety of L2 proficiency in a more appropriate way. Table 2 shows JLs' proficiency levels in the present study. Over 80% of the JLs in the data are categorized in A2 or B1_1, which indicates that many of them are mostly in the pre-intermediate or intermediate level. Thus, this study sheds some light on an early stage of the acquisition of the verb *get*.

	A2	B1_1	B1_2	B2+
JLs	38.5	44.8	12.3	4.5

Table 2: *Ratios of JLs at the Four Proficiency Levels (%)*

NSs in the data consist of 100 college students and 100 non-college students. They wrote essays in exactly the same writing conditions as EFL learners did, which makes the comparison more powerful and reliable. Their nationalities include the Unites States (57%), the United Kingdom (14%), Australia (8.5%), and New Zealand (6.5%).

3.2 Procedure

In the first step of the analysis, I extracted and counted all the instances of the verb *get* in the data. I used a concordancing program, *AntConc 3.3.5m* for this process (Anthony, 2013). By examining the overall frequencies of *get*, I analyzed whether JLs overused or underused the verb *get*. In the second step, I classified every instance into eight groups depending on their grammatical patterns, counted the frequency of each group, and examined the differences between JLs' use and NSs' use. Frequency differences across the samples were tested by means of the chi-square test, with 5% ($p < .05$) as the critical level of significance. The frequencies of *get* were compared with the total number of words in each sample. This method corresponded to the analysis adopted

in the previous studies, Altenberg and Granger (2001) and Mochizuki (2007). When the observed values are less than five, the chi-square test is not appropriate. In such cases, I used Fisher's exact test to compare the two sets of data.

4. Results and Discussion

4.1 Overall Frequency of the Verb *Get*

First, overall frequency of the verb *get* is compared between JLs and NSs. Table 3 shows the frequencies of the verb *get* in JL and NS writing and chi-square values between the two groups. To compare frequencies in different-sized data, the normalized frequency is provided with the raw frequency.

	JLs	NSs	χ^2
get	714	271	16.44 ($p < .001$) **overuse**
normalized *get* (per 100,000 words)	398.8	299.1	N/A

Table 3: *Frequency of the Verb* Get *in JL and NS Writing*

The result shows that there are nearly three times as many occurrences of the verb *get* in JL writing as in NS writing, and that JLs significantly overuse the verb *get* than NSs do ($p < .001$). This tendency of overuse is in accordance with the JLs' use of the verb *make* in Mochizuki (2007). Compared with other high-frequency verbs, *get* is relatively lower in frequency in native written English, and thus little emphasis has been placed on the use of it. From JLs' point of view, however, this study reconfirms the observation that the verb *get* is clearly one of basic verbs JLs heavily rely on when producing English sentences.

4.2 Major Uses of the Verb *Get*

The verb *get* has many different meanings and grammatical patterns. Based on the categories provided in Biber, Johansson, Leech, Conrad, and Finegan (1999), the major uses of the verb *get* are classified into eight categories. Table 4 lists the coding categories and illustrates each of them with some examples found in the data.

Grammatical patterns	Examples
1) *get*+Adverbial	*get in school*
2) *get*+Adjective	*get sleepy, get happier*
3) *get*+Noun/NP	*get a part-time job, get enough sleep*
4) Ditransitive *get*	*get them a job*
5) Causative *get*	*get them ready, get them to quit smoking*
6) have got to do	*have got to know the basics*
7) *get* to do	*get to know the risk*
8) Phrasal/prepositional *get*	*get used to, get along with, get rid of, get by*

Table 4: *Coding Categories: Major Uses of the Verb* Get

I coded every instance of the verb *get* into the eight categories above. Table 5 provides the results of the classification and chi-square values between the frequency in JL/NS writing and the total number of words in each sample. As is clear from Table 5, there is not much difference in the rank orders of eight constructions between JLs and NSs. *Get*+NP is most common, followed by *get*+Adjective and phrasal/prepositional *get*. However, statistical test results show significant differences in several constructions. JLs significantly overuse the *get*+Noun/NP construction ($p < .001$), and they underuse the *get*+Adjective construction and "have got to do" construction ($p < .01$). Among the three constructions, *get*+Noun/NP displays the most striking difference between JLs and NSs. The frequency of this construction in JL writing is more than three times that seen in NS writing. The following sections look into this most frequent category, the *get*+Noun/NP construction, in more detail.

	JLs	NSs	χ^2 / p values
1) *get*+Adverbial	2	3	$p = .34$, ns (Fisher's exact test)
2) *get*+Adjective	48	43	7.60 ($p < .01$) **underuse**
3) *get*+Noun/NP	563	152	48.96 ($p < .001$) **overuse**
4) Ditransitive *get*	1	2	$p = .26$, ns (Fisher's exact test)
5) Causative *get*	8	9	$p = .12$, ns (Fisher's exact test)
6) have got to do	1	7	$p < .01$, **underuse** (Fisher's exact test)
7) *get* to do	16	9	0.06 ($p = .80$, ns)
8) Phrasal/ prepositional *get*	71	44	1.12 ($p = .29$, ns)
Frequency of *get*	714	271	16.44 ($p < .001$) **overuse**
Number of words	179,042	90613	

Table 5: *Uses of the Verb* Get *in JL and NS Writing*

4.3 *Get*+Noun/NP Constructions

When JLs use the verb *get*, they heavily rely on using the verb in the *get*+Noun/NP construction. Then, what types of nouns are used by JLs and NSs? Do they use this construction in a similar or different way? Table 6 shows the list of 15 nouns that are commonly used with the verb *get*[2]. It shows that some nouns (i.e., *job*, *skill*, *experience*) are common in both JL and NS writing. Other nouns, however, behave in a quite different manner. The most prominent example is *get+money*. This combination is frequently used by JLs but does not appear in the list for NSs.

It is also important to note that NSs use the verb *get* with quite limited types of nouns such as *job*, *experience*, *grade*, or *skill*, whereas JLs tend to connect the verb with a wide range of nouns. In fact, the type frequency of NSs' *get*+Noun/NP construction is only 47, 35 of which appear only once in the data, while JLs produce 101 types of this construction. One plausible explanation for this tendency might be that JLs still rely on Sinclair (1991)'s open-choice principle to make such atypical combinations. JLs seem to produce all the possible combinations that the generalized meaning of *get* can take. This tendency also

	JLs		NSs	
1	*money*	209	*job*	65
2	*job*	65	*experience*	13
3	*skill*	27	*grade*	10
4	*cancer*	25	*skill*	7
5	*experience*	23	*degree*	5
6	*thing*	15	*cancer*	4
7	*friend*	9	*it*	3
8	*chance*	8	*score*	3
9	*it*	8	*work*	3
10	*knowledge*	7	*career*	2
11	*work*	7	*everything*	2
12	*disease*	6	*sleep*	2
13	*license*	6	*air*	1
14	*opportunity*	6	*average*	1
15	*pleasure*	6	*bee*	1

Table 6: *Top 15 Direct Objects of the Verb* Get

implies a difference between JLs' and NSs' acquisition of the generalized meaning of the verb *get*. NSs construct some generalized meanings of the verb *get* by abstracting them from collocations based on concrete contexts or settings. Instead, as JLs are more familiar with generalized meanings, which are more emphasized in teaching and testing in classrooms, they lack NSs' collocational knowledge of the verb and its selectional restrictions.

As for the top 10 nouns in JL writing, I counted the number of NSs' uses and compared the frequencies by conducting the chi-square test on the two groups. As shown in Table 7, JLs significantly overuse some combinations such as *get+money/thing/friend* and underuse *get+job*. It should be noted that the combinations JLs overuse, except for *get+cancer*, are never produced by NSs in the data. On the contrary, the most common combination among NSs, *get+job*, is significantly underused by JLs. This result indicates that there is a striking difference in the way of using the verb *get* in the *get*+Noun/NP construction between JLs and NSs. Previous studies have revealed that some atypical combinations are the result of L1 transfer. However, *get+friend* is not likely to be the case, because *tomodachi wo tsukuru*, meaning "make friends," is more common than *tomodachi wo eru*, meaning "get friends," in Japanese.

	JLs	NSs	χ^2 / p values
get+money	209	0	$p < .001$, **overuse** (Fisher's exact test)
get+job	65	65	15.67 ($p < .001$) **underuse**
get+skill	27	7	2.58 ($p = .11$, ns)
get+cancer	25	4	$p < .05$, **overuse** (Fisher's exact test)
get+experience	23	13	0.10 ($p = .75$, ns)
get+thing	15	0	$p < .01$, **overuse** (Fisher's exact test)
get+friend	9	0	$p < .05$, **overuse** (Fisher's exact test)
get+chance	8	0	$p = .06$, ns (Fisher's exact test)
get+it	8	3	$p = .76$, ns (Fisher's exact test)
get+knowledge	7	3	$p = 1.00$, ns (Fisher's exact test)

Table 7: *Uses of* Get+*Noun/NP Construction in JL and NS Writing*

The following sentences are examples of JLs' *get+money/thing/friend* constructions taken from the data. Although these

sentences cannot be regarded as ungrammatical, they would give an awkward and inappropriate impression to readers, and it would be easy to infer that the sentences do not result from nativelike selection of expressions.

(1) Examples of the *get+money* construction
   a. … and I would like to understand how hard it is to ***get some money***. (JPN_PTJ_070_A2)
   b. Students may *get money* from parents, but I think that they should ***get money*** by themselves because … (JPN_PTJ_162_A2)
   c. So, they must know the difficulty of ***getting money***. (JPN_PTJ_251_B1_1)

(2) Examples of the *get+friend* construction
   a. We want to ***get good friends***. (JPN_PTJ_228_A2)
   b. And you can ***get a close friend*** in part-time job. (JPN_PTJ_107_B1_1)
   c. I think a part-time job is a good opportunity to ***get new friends***. (JPN_PTJ_015_B1_2)

(3) Examples of the *get+thing* construction
   a. A part-time job has ***the things that I do not get*** in the college life. (JPN_PTJ_206_A2)
   b. They can not only earn money but also can ***get some good things*** from part time jobs. (JPN_PTJ_359_B1_1)
   c. … college students can ***get important things*** from not only studies but also … (JPN_PTJ_279_B1_2)

As for the combinations of *get+money* and *get+friend*, it should be better to replace them with *make money/earn money* or *make friends*. Given that NSs never produce such combinations, it is unlikely that JLs have heard or read them in the input. One explanation for this use is that they may be using the verb *get* or the phrase *get+money* or *get+friend* as lexical teddy bears. As it is uncertain whether JLs consider these combinations awkward, they might be employing the strategy to stick to familiar words so that they could avoid making errors and stay within their safety zone.

Another explanation is the nature of input. Since most JLs

learn English as a foreign language, the input for them tends mainly to be teaching materials used in middle and high schools in Japan. JLs' overuse of *get+money/thing/friend* can be explained by the idea that the language presented in textbooks fails to capture the essence of real use and fails to help them expand their vocabulary usefully in a wide range of contexts including academic, business, or working settings. This view is corroborated by the fact that JLs overuse not only the verb *get* but also *money* in general ($\chi^2 = 284.25$, $p < .001$). The token frequency of *money* is only 212 in NS writing but is 1354 in JL writing. On the one hand, NSs use a wide range of expressions including *cash, expenses*, *fees*, *getting paid, income, internship*, *scholarship*, or *tuition*, when referring to events involving money exchange or employment. On the other hand, JLs tend to cling on to the word *money* and use similar phrases repeatedly. For example, (4) is an example of JL writing. Here, the learner used the word *money* 4 times within five sentences. In a similar manner, some JLs at A2 level in the data use *money* 7 to 10 times in their essays.

(4) … The understanding of difficulty of ***getting some money*** will make them to thank their parents. Finally college students ***need a lot of money***. For example they ***need money*** when they go lunch with their club members or they go shopping. They have to ***make money*** by themselves for their enjoyment. Therefore I think a part-time job is important for them…. (JPN_PTJ_070_A2)

What is more complicated, the learner above used *get+money* even though he or she knew another conventional combination *make+money*. This use implies that JLs have problems with collocations involving not only high-frequency verbs but also other basic words.

5. Conclusion

Previous studies report that EFL learners have great difficulty with high-frequency verbs. The present study also confirms the complexity of the use of the verb *get* in JL writing. JLs tend to overuse the verb *get* overall, which can be attributed to the fact that JLs heavily rely on the *get*+Noun/NP construction. At the

same time, they significantly underuse other constructions (i.e., *get*+Adjective, "have got to do" construction). Looking into the *get*+Noun/NP construction in more detail, the present study demonstrates that they use many atypical combinations (e.g., *get+money*, , get+*friend*, *get+thing*) by connecting the verb *get* with a variety of nouns. This finding suggests that JLs lack collocational knowledge of the verb *get* and tend to rely on the open-choice principle when they produce English sentences. This tendency is unfortunate because they must have encountered the verb early in classrooms.

These results have interesting pedagogical implications. As many of the atypical combinations derive from a deficient knowledge of collocations, it is necessary to raise awareness of the wide variety of structures high-frequency verbs can take. There are several approaches to achieve this pedagogical purpose. One approach is teaching verbs together with their collocational patterns (Granger, 2011). Some researchers are currently engaged in listing such collocations in a manner comparable with the Academic Word List (Coxhead, 2000; Durrant, 2009; Simpson-Vlach & Ellis, 2010), so these lists might be useful for instructors and students. Another approach is data-driven learning. Data-driven learning (DDL) is defined as the use of computer-generated concordances to get students to explore restrictions of patterns in the target language (Johns, 1991). This method draws learners' attention to collocations inductively. For example, instructors can provide concordance lines of a high-frequency verb written by JLs and NSs and ask students to find and discuss differences between JLs and NSs.

The nature of input should also be taken into consideration in explaining the differences. JLs lack collocations involving not only the verb *get* but also other basic words. To expand their collocational knowledge, instructors can provide learners with specific discourses and tasks based on the discourse. This problem might show different aspects of learners from different ESL countries. Since previous corpus-based studies largely ignore the educational context in each country and assume that the findings should be applicable for learners in general (Tono, 2009), it is necessary to compare the use among learners with different L1s. Fortunately, such learner data is available in the ICNALE. Further research can contribute to a better understanding to give

contrastive descriptions of learners' use of the verb *get*.

Notes

1. Instead of "collocations," various terms are used for such combinations, including "idioms," "formulaic language," "fixed (fossilized) expressions," or "multi-word units" (Wray, 2002).
2. Direct objects of the verb *get* often appear as the combinations of an adjective and a noun. In order to capture a broader and clearer picture of the use of the verb *get*, I did not take the adjectives into consideration when classifying the instances. For example, *get much money*, *get enough money*, and *get some money* are all classified as instances of the identical *get*+Noun/NP construction.

References

Allerton, D. J. (1984). Three (or four) levels of co-occurrence restriction. *Lingua*, *63*, 17–40.

Altenberg, B., & Granger, S. (2001). The grammatical and lexical patterning of MAKE in native and non-native student writing. *Applied Linguistics*, *22*(2), 173–195.

Anthony, L. (2013). AntConc (Version 3.3.5m) [Computer Software]. Tokyo, Japan: Waseda University. Available from http://www.antlab.sci.waseda.ac.jp/

Biber, D., & Reppen, R. (1998). Comparing native and learner perspectives on English grammar: A study of complement clauses. In S. Granger (Ed.), *Learner English on computer* (pp. 145–158). London: Longman.

Biber, D., Johansson, S., Leech, G., Conrad, D., & Finegan, E. (1999). *Longman grammar of spoken and written English*. Harlow: Pearson.

Biber, D., Conrad, S., & Cortés, V. (2004). *If you look at…*: Lexical bundles in university teaching and textbooks. *Applies Linguistics*, *25*(3), 371–405.

Borin, L., & Prütz, K. (2004). New wine in old skins? A corpus investigation of L1 syntactic transfer in learner language. In G. Aston, S. Bernardini, & D. Stewart (Eds.), *Corpora and language learners* (pp. 67–87). Amsterdam, The Netherlands: John Benjamins.

Coxhead, A. (2000). A new academic wordlist. *TESOL Quartely*, *34*(2), 213–238.

Durrant, P. (2009). Investigating the viability of a collocation list for students of English for academic purposes. *Journal of English for Specific Purposes*, *28*(3), 157–169.

Granger, S. (1996a). From CA to CIA and back: An integrated approach to computerized bilingual and learner corpora. In Aijmer K., Altenberg B.,

& M. Johansson M. (Eds.), *Language in contrast: Text-based cross-linguistic studies* (pp. 37–51). Lund: Lund University Press.

Granger, S. (1996b). Romance words in English: From history to pedagogy. In J. Svartvik (Ed.), *Words: Proceedings of an international symposium* (pp. 105–121). Stockholm: Almqvist and Wiksell International.

Granger, S. (Ed.). (1998). *Learner English on computer*. London, UK: Longman.

Granger, S. (2011). From phraseology to pedagogy: Challenges and prospects. In T. Herbst, P. Uhrig & S. Schller (Eds.), *Chunks in the description of language: A tribute to John Sinclair* (pp. 123–146). Berlin & New York: Mouton de Gruyter.

Groom, N. (2005). Pattern and meaning across genres and disciplines: An exploratory study. *Journal of English for Academic Purposes*, *4*(3), 257–277.

Hasselgren, A. (1994). Lexical teddy bears and advanced learners: A study into the ways Norwegian students cope with English vocabulary. *International Journal of Applied Linguistics*, *4*, 237–260.

Howarth, P. (1988). The phraseology of learners' academic writing. In A. P. Cowie (Ed.), *Phraseology: Theory, analysis, and applications* (pp. 161–186). Oxford: Oxford University Press.

Ishikawa, S. (2013). The ICNALE and sophisticated contrastive interlanguage analysis of Asian learners of English. In S. Ishikawa (Ed.), *Learner corpus studies in Asia and the world Vol. 1*, (pp. 91–118). Kobe: Kobe University.

Jackendoff, R. (1997). Twistin' the night away. *Language, 73*(3), 534–559.

Johns, T. (1991). From printout to handout: Grammar and vocabulary teaching in the context of data-driven learning. *English Language Research Journal, 4*, 27–45.

Källkvist, M. (1999). *Form-class and task-type effects in learner English: A study of advanced Swedish learners.* Lund: Lund University Press.

Lennon, P. (1996). Getting "easy" verbs wrong at the advanced level. *International Review of Applied Linguistics in Language Teaching*, *34*, 23–36.

Marco, M. J. L. (2011). Exploring atypical verb+noun combinations in learner technical writing. *International Journal of English Studies*, *11*(2), 77–95.

Mochizuki, M. (2007). Nihonjin daigakusei no EFL gakushusha corpus ni mirareru make no shiyo. *Kansai Daigaku Gaikokugo Kyoiku Kenkyu*, *14*, 31–45. [The uses of MAKE in EFL learner corpus of Japanese university students].

Nattinger, J., & DeCarrico, J. (1992). *Lexical phrases and language teaching*. Oxford: Oxford University Press.

Nesselhauf, N. (2004). Learner corpora and their potential for language teaching. In J. Sinclair (Ed.), *How to use corpora in language teaching* (pp. 125–152). Amsterdam: John Benjamins.

Pawley, A., & Syder, F. H. (1983). Two puzzles for linguistic theory: Native-like selection and nativelike fluency. In J. C. Richards & R. W. Schmidt

(Eds.), *Language and communication* (pp. 191–225). London: Longman.
Simpson-Vlach, R. & Ellis, N. C. (2010). An academic formulas list: New methods in phraseology research. *Applied Linguistics*, *31*(4), 487–512.
Sinclair, J. (1991). *Corpus, concordance, collocation*. Oxford: Oxford University Press.
Stubbs, M. (2002). *Words and phrases: Corpus studies of lexical semantics.* Oxford: Blackwell Publishing Ltd.
Tono, Y. (2009). Integrating learner corpus analysis into a probabilistic model of second language acquisition. In P. Baker (Ed.), *Contemporary corpus linguistics* (pp. 184–203). London, UK: Continuum.
Trask, R. (1993). *A dictionary of grammatical terms in linguistics*. London: Routledge.
Viberg, Å. (1996). Cross-linguistic lexicology: The case of English *go* and Swedish *gå*. In K. Aijmer, B. Altenberg & M. Johansson (Eds.), *Language in contrast* (pp. 151–182). Lund: Lund University Press.
Warren, B. (2005). A model of idiomaticity. *Nordic Journal of English Studies*, *4*(1), 35–54.
Wray, A. (2002). *Formulaic language and the lexicon.* Cambridge: Cambridge University Press.

Appendix: Instruction Sheet Given to Learners (Ishikawa, 2013)

Do you agree or disagree with the following statements? Use reasons and details to support your opinion.

(Topic A) It is important for college students to have a part-time job.
(Topic B) Smoking should be completely banned at all the restaurants in the country.

1. Clarify your opinions and show the reasons and some examples.
2. You can use 20 to 40 minutes for each essay. This means that you have 40 to 80 minutes to complete two essays. Do not finish too early or spend too much time.
3. You must use MS Word or a similar word processor.
4. Do not use dictionaries or other reference tools.
5. Do not plagiarize anyone else's essays.
6. The length of your single essay should be from 200 to 300 WORDS (not letters). Too short or too long essays cannot be accepted. You can check the length of your essay using the word count function of MS Word.
7. You must run spell check before completing your writing.