THE JOURNAL OF ASIA TEFL
Vol. 16, No. 1, Spring 2019, 279-293
http://dx.doi.org/10.18823/asiatefl.2019.16.1.18.279

The Journal of Asia TEFL

http://journal.asiatefl.org/

e-ISSN 2466-1511

A Corpus-Based Study on "Delexical Verb + Noun" Collocations Made by Korean Learners of English

Wonkyung Choi
The University of Birmingham, UK

Collocation is one of the essential elements that must be acquired in the second language classroom to reach an advanced level of English proficiency. However, L2 learners of English tend to have difficulty in using native-like collocations. It has been reported that "delexical verb + noun" collocations are particularly difficult for L2 learners from various L1 backgrounds, owing to the fact that delexical verbs are characterized by little or depleted meaning. The present study uses corpus-based analyses to investigate how Korean learners of English use delexical collocations. It compares a learner corpus comprising essays written by Korean university students with a native speaker corpus as a control. The findings of a quantitative analysis indicate that Korean learners significantly underuse delexical collocations and that it would be hard for them to reach a native-like command with regard to this language element. The findings of a qualitative analysis show that Korean learners make quite a few collocational errors with delexical verbs, and that these errors seem to be caused by both interlingual and intralingual influences. These results offer meaningful pedagogic implications for language teachers and researchers in the field of English language teaching.

Keywords: delexical verb, collocation, learner corpus research, Korean learners of English

Introduction

Delexical verbs[1] can be defined simply as "verbs with little meaning" (Sinclair, 1990, p. 147). They are very common transitive verbs that adopt a noun as an object to show that someone performs an action, e.g., *have a look*, *make a decision*, and *take a bath*. For example, in the delexical collocation *make a mistake*, the verb *make* would be referred to as a delexical verb and the noun *mistake* as an object. In verb + noun collocations, delexical verbs do not retain their specific literal sense, while the noun (object) determines the meaning. Hence, these verbs look semantically empty or depleted, which renders the term "delexical verb." The delexical verb + noun collocations[2] are quite frequently used in place of a simple verbal form (e.g., *make a mistake* for *mistake*), presumably because these constructions allow different conceptualizations of a given situation. As Allan (1998) explained:

[1] Some researchers use the term "light verb" rather than "delexical verb." Jespersen (1965, p. 117) coined the term light verb for verbs that must take nouns, such as *look*, *bite*, as in *have a look*, *take a bite*. However, sometimes the two terms are used interchangeably (e.g., Liu, 2010; Nesselhauf, 2005). In the present study, the meaning of delexical verb encompasses all these types of transitive verbs.

[2] In the present study, the phrase "delexical verb + noun collocation" refers to the construction in which the delexical verb adopts a noun as an object, and it is used interchangeably with "delexical construction" and "delexical collocation."

> When a speaker uses a simple verbal form such as *look*, the focus is on the action; whereas when using the nominal form of the same verb in a delexical structure such as *have a look*, the speaker is naming an event and in so doing, delimiting an activity or state which in its verbal form lacks temporal boundaries or a sense of completeness (p. 1).

It has been argued that all languages have these kinds of delexical verbs, although their usage and collocates might differ (Altenberg & Granger, 2001). In English, the high-frequency verbs such as *do*, *get*, *give*, *have*, *make,* and *take* are often used as delexical verbs (Howarth, 1996). Note that these verbs are used not only in a delexical manner, but also in many other ways with various senses and usages. For example, Altenberg and Granger (2001) presented at least eight major usages of the verb *make*. The polysemous nature and the complexity in usage of these high-frequency verbs may present obstacles to L2 learners. Indeed, it has been reported that many L2 learners from various L1 backgrounds have difficulty in the acquisition of delexical verbs (Babanoglu, 2014; Chi, Wong, & Wong, 1994; Jukneviciene, 2008; Kittigosin & Phoocharoensil, 2015; Lennon, 1996; Nesselhauf, 2005; Wang, 2016; Yan, 2006).

The present study investigates delexical verb + noun collocations in a learner corpus that consists of essays produced by Korean university students. Using a native speaker corpus as a norm for comparison, it discovers how Korean learners under- or overuse delexical constructions with four high-frequency verbs (*make*, *take*, *give*, and *get*), in order to find out whether they develop a native-like level of command in the use of this feature. In addition, this study looks into deviant delexical collocations with the four verbs and speculates as to their possible causes. Researchers have investigated the use of high-frequency verbs in learner corpora or in teaching materials. However, to the best of my knowledge, to date there has been no study that examines Korean learners' collocational errors associated with delexical verbs. Therefore, the present study is expected to provide valuable implications for researchers and teachers in the English classroom.

Theoretical Basis of Delexical Collocation

Collocation in English Language Teaching (ELT)

This study focuses on a specific type of verb-noun collocation. Therefore, it is necessary to clarify the concept of collocation, because the term "collocation" is widely used across previous research, but with slightly different meanings (Walker, 2008). The Cambridge English Dictionary[3] defines collocation as "the combination of words formed when two or more words are often used together." According to Nation (2001, p. 317), however, collocation refers to "items which frequently occur together and have some degree of semantic unpredictability." Furthermore, a number of studies that have investigated collocation have used alternative terms such as prefabricated units, lexical bundles, (lexical) chunks, phraseological units, multi-word units (MWU), and formulaic sequences. Simply put, collocation can be defined as "some kind of syntagmatic relationship between words" (Nesselhauf, 2005, p. 11).

Collocation has been a fundamental issue in the research and practice of ELT since the 1990s. Language researchers and teachers have noted the significance of collocations in language classrooms, as it is acknowledged that L2 learners acquire language as multiple units (or chunks) rather than as single words. For example, Lewis (1993) presented the Lexical Approach, in which collocations should be treated as crucial in the ESL/EFL classroom, based on the belief that "a central element of language teaching is raising students' awareness of, and developing their ability to 'chunk' language successfully" (p. vi). Hill (2000) presented seven typical patterns of collocations (e.g., adjective + noun, noun + noun, verb + adjective + noun) to which students should pay particular attention. Above all, collocation is one

[3] It can be accessed at the URL: http://dictionary.cambridge.org/dictionary/english.

of the major elements that should be taught to L2 learners, and the importance of teaching collocations in the ESL/EFL classroom cannot be stressed too highly.

L2 Acquisition of Delexical Collocation

Many researchers have investigated L2 acquisition of collocations through analyses of learners' errors (Hong, Rahim, Hua, & Salehuddin, 2011; Laufer & Waldman, 2011; Nesselhauf, 2005; Zheng & Xiao, 2015), which suggest that L2 learners are likely to have difficulties in producing native-like collocations. In particular, several researchers have reported L2 learners' errors with the high-frequency verbs (e.g., Altenberg & Granger, 2001; Gouverneur, 2008; Howarth, 1996; Lennon, 1996; Nesselhauf, 2005). These high-frequency verbs, such as *do*, *have*, *make,* and *take*, had long been considered to be "easy" words for L2 learners in the early stages of L2 acquisition (Lennon, 1996; Wang, 2016), probably because L2 learners tend to focus only on core literal senses of the verbs and rather neglect extended meanings. However, as Wang (2016, p. 1) argued, the high-frequency verbs may be "a stumbling block to native-like proficiency in language production." This might be because, when the high-frequency verbs are used in delexical collocations, they have little or no independent meaning (Wang, 2016).

For example, Nesselhauf (2005) reported that German learners of English made various kinds of verb-noun collocations with delexical verbs, and argued that one possible cause of these errors would be L1 (German) transfer. Jukneviciene (2008) investigated collocations with five high-frequency verbs (*have*, *do*, *make*, *take,* and *give*) produced by Lithuanian learners of English. Chi et al. (1994) examined collocational errors of delexical verbs that occurred in the Hong Kong University of Science & Technology (HKUST) learner corpus.

A number of researchers and teachers in Korea have also noted the significance of delexical constructions in ELT. However, the majority of previous studies in Korea have analyzed the distribution of the usages of high-frequency verbs in the textbooks or teaching materials rather than how Korean EFL learners produce delexical constructions (Choi, 2012; Jang, 2008; Kang, 2014; Kwon, 2012; Seo, 2011; Shin, 2012). While a few studies have investigated Korean learners' uses of English delexical verbs (Kim, 2002; Lee & Na, 2015), regrettably, these replication studies of Altenberg and Granger (2001) looked only at the use of a single verb, *make*. More recently, Ha and Nam (2018) conducted learner corpus research in which they analyzed three "light" verbs (*make*, *take*, and *get*) used in a Korean learner corpus in comparison with a native speaker corpus. However, rather than focusing on the delexical usage of the verbs, they also included occurrences of the given verbs with literal senses. In contrast, the investigation in the present study is expected to shed light on how Korean EFL learners acquire and use delexical collocations.

Possible Causes of L2 Learners' Deviant Collocations (Collocational Errors)

Several researchers have examined the possible causes of L2 learners' deviant delexical collocations. Their findings indicate that these errors have two main sources: interlingual and intralingual influences.

First, previous research has shown that interlingual influence (or L1 transfer), which refers to the effect of L2 learners' mother tongue, seems to play a significant role in producing such errors. For example, Nesselhauf (2003, 2005) reported that some deviant delexical collocations made by German learners of English appear to be affected by their L1 (German). Kittigosin and Phoocharoensil (2015) conducted a translation test that was designed to encourage 80 Thai learners of English to produce delexical constructions. They found that more than 50% of the errors were related to interlingual influence. Vrbinc (2004) and Jukneviciene (2008) found a similar L1 effect with Slovene and Lithuanian learners of English respectively. Therefore, interlingual influence may be one of the major factors that affect deviant delexical collocations.

Second, other studies have revealed an intralingual influence on deviant delexical collocations. For example, Liu (2010) investigated collocational errors with three delexical verbs (*have*, *make*, and *take*)

made by 200 Taiwanese university students. Liu found that Taiwanese learners sometimes selected incorrect delexical verbs (e.g. **have a mistake* for *make a mistake*, **take a statement* for *make a statement*, **make a photograph* for *take a photograph*), and argued that these errors derived from the overgeneralization of delexical verbs, i.e. intralingual influence. Gonzalez Alvarez and Doval Suarez (2010) reported similar kinds of deviant delexical collocations made by Spanish learners of English. These errors appear to be caused by L2 learners' lack of awareness of the delexical constructions. As Chi et al. (1994, p. 162) indicated, "As such verbs carry no significant meaning, it is likely that students will choose the wrong verb-noun collocation unless they have previously learned it as a chunk."

Therefore, the present study investigates whether Korean learners of English are also affected by both interlingual and intralingual influences when producing delexical collocations. To do so, it conducts a qualitative analysis and illustrates some examples of the errors affected by each kind of influence.

Research Design

Research Questions

Based on the review of literature on L2 learners' use of delexical collocations, this study addresses the following research questions (RQs):

1. Does the frequency of delexical verb + noun collocations in a learner corpus differ from that in a native English corpus?
2. Does the frequency of delexical verb + noun collocations in a learner corpus differ according to Korean learners' English proficiency levels?
3. Can interlingual and intralingual influences be identified as the possible causes of the deviant delexical verb + noun collocations that occur in a learner corpus?

In order to answer research question (RQ) 1, I compared the frequencies of delexical verb + noun collocations between a Korean learner corpus, the Yonsei English Learner Corpus (YELC) (Rhee & Jung, 2014), and a native English corpus, the Louvain Corpus of Native English Essays (LOCNESS).

To address RQ 2, I compared the frequencies of delexical verb + noun collocations in each sub-corpus of the YELC, each of which is divided into three groups according to learners' English proficiency level (see Table 1).

In order to answer RQ 3, I investigated the concordance lines in more detail through a qualitative analysis. I used BNCweb to find collocational errors of delexical verb + noun combinations in the YELC, and speculated as to the possible causes of these errors: interlingual or intralingual influence.

A Korean Learner Corpus and a Native English Corpus

This study analyzed a Korean learner corpus, the YELC. Compiled in 2011, it consists of essays written by freshmen at Yonsei University, Seoul, Korea. The essays were created from the Yonsei English Placement Test that the freshmen had to take to be used to diagnose their English proficiency. This test consists of two parts where they are required to write narrative and argumentative essays (Rhee & Jung, 2014). One of the merits of this learner corpus is that it contains a very large dataset (approximately 1 million words). Considering that the sub-corpora of the International Corpus of Learner English (ICLE) (Granger, Dagneaux, Meunier, & Paquot, 2009), acknowledged as a prototypical learner corpus, comprise 100,000 to 200,000 words, the YELC is five times larger. Therefore, it can be expected that the YELC will contain a considerable number of delexical verb + noun collocations. Another merit of the YELC is that it is sub-divided into nine proficiency levels according to the refined version of CEFR (Common European Framework of Reference for Languages) writing scales, as shown in Table 1 (Rhee

& Jung, 2014). Consequently, the present study can compare the frequencies of the collocations within each proficiency level.

TABLE 1
The Distribution of the YELC by Grades

Level	Grade	Number of texts	Number of words
Low	A1	82	3,056
	A1+	370	36,009
	A2	1,368	195,473
	Sum	1,820	234,538
Intermediate	B1	2,346	391,463
	B1+	1,410	263,470
	B2	756	146,843
	Sum	4,512	801,776
Advanced	B2+	162	33,250
	C1	74	15,434
	C2	4	829
	Sum	240	49,513
	Total	6,572	1,085,827

In order to compare EFL use of collocations with native English use, it is necessary to have a native speaker control corpus (Altenberg & Granger, 2001). The present study used the LOCNESS as a norm for comparison with the YELC. Although the LOCNESS is somewhat smaller (324,304 words) than the YELC, it is arguably a relevant corpus to use as a norm for native English because, like the YELC, it consists of argumentative essays written by university students, in this case native English speakers from the US and UK. Given the high degree of commonality between the two corpora, the LOCNESS[4] is relevant as a norm to analyze learner data from the YELC.

Analyses to Detect Delexical Collocations and the Collocational Errors with Delexical verbs

In order to investigate how Korean learners use delexical collocations in comparison with native speakers, this study adopts both quantitative and qualitative analyses. With regard to the former, a corpus software program is utilized to detect the frequency of delexical collocations that occur in each corpus. This analysis shows whether Korean learners under- or overuse delexical collocations compared with native speakers. Then, statistical measurements are conducted to find out whether or not the difference between Korean learners and native speakers is significant. For the qualitative analysis, a workable number of concordance lines in a learner corpus are thoroughly investigated to find deviant delexical collocations (or collocational errors). In each analysis, it is necessary to use clear criteria to determine corresponding delexical collocations and to find deviant collocations.

First, identifying whether the verbs in collocations have a delexical meaning is not simple, because it is a matter of the degree to which the verbs retain core literal senses or become depleted (Wang, 2016). Previous research has used two main methods to detect delexical verb + noun collocations. In the first method, researchers look into concordance lines that contain target high-frequency verbs one by one, in order to find relevant collocations in which the verbs are used delexically. For example, Howarth (1996) presented a series of tests that might be useful to identify delexical verb + noun collocations. However, this method would be time-consuming, especially when a large corpus dataset is analyzed as in the present study. The second method is to utilize a predetermined list to detect relevant collocations. Some researchers have used examples of delexical collocations presented in English dictionaries (e.g., Jang, 2008; Kang, 2014; Seo, 2011). However, English dictionaries provide only a limited number of delexical

[4] More information about the LOCNESS is available at https://uclouvain.be/en/research-institutes/ilc/cecl/locness.html.

collocations. Therefore, the present study uses Wang's (2016) list of noun collocates of the four high-frequency verbs, determined according to statistical measures from the British National Corpus (BNC). According to Wang (2016, p. 16), the list includes "the overwhelming majority of noun collocates from the BNC … to identify delexical collocations." Hence, the list can almost certainly detect most, if not all, delexical verb + noun collocations in the given corpora. Although it may also contain some noun collocates that are not related to delexical verb + noun collocations, the occurrence of irrelevant collocates is undoubtedly too small to affect the main results of the present study.

Second, this study uses a qualitative analysis to look into deviant delexical collocations in an attempt to speculate as to the possible causes of deviation. Again, it might be difficult, though not impossible, to determine which delexical collocations should be considered to be "deviant" or collocational errors. The errors should be perceived as on a continuum of deviation, rather than on one side of a dichotomy between correct and incorrect usage. In order to determine collocational errors, some previous studies have used native speakers' intuition (e.g., Altenberg & Granger, 2001; Lee & Na, 2015; Nesselhauf, 2005) or have referred to English dictionaries (e.g., Chi et al., 1994). However, the present study employs a reference corpus, the BNCweb, which includes many more delexical collocations than any English dictionary. Consequently, if the frequency of a delexical verb + noun collocation is significantly low in the BNCweb, then one can infer that native speakers never or rarely use this collocation, and it can therefore be regarded as a deviant collocation (or a collocational error) newly created by L2 learners. Since the present study determines the degree of deviation according to a statistical measure and cut-off point commonly used in corpus linguistics, the results as to whether the delexical collocations are deviant or not can be considered reliable.

Data Analysis Procedure

This study conducted four steps of data analysis to investigate delexical verb + noun collocations, as shown in Figure 1.

In Step 1, I selected four delexical verbs (*make*, *take*, *give*, and *get*) for analysis, based mainly on their frequency in language use. Previous studies have investigated the same four delexical verbs (e.g., Altenberg & Granger, 2001; Chi et al., 1994; Jukneviciene, 2008; Kim, 2002; Kwon, 2012; Liu, 2010). Note that *have* and *do*, two high-frequency verbs that have often been analyzed in previous research, were excluded from the present study. This is because these two verbs can also be used as auxiliary verbs, which means that in the course of data analysis the corpus software program would retrieve too many irrelevant concordance lines. These irrelevant occurrences would then have to be deleted manually, which might be laborious and time-consuming.

In Step 2, I prepared the corpus data (i.e., the YELC and the LOCNESS) for analysis, and the list of noun collocates for detecting relevant delexical verb + noun collocations. Since there are many instances in which the high-frequency verbs are not used as delexical verbs (e.g., the verb *make* can be used with the meaning "produce"), it was necessary to obtain the frequencies of delexical verb + noun collocations relevant to the present study. As mentioned earlier, this study employs Wang's (2016) list of noun collocates of the four verbs. Wang (2016) provided the first 100 noun collocates ranked in the BNC according to two statistical measures, log-likelihood (LL) and mutual information (MI), for two forms (present and past) of each verb. Consequently, each delexical verb has 400 noun collocates, which overlap between verbs. The present study uses only the different types of noun collocates, which is around 200 collocates for each verb.

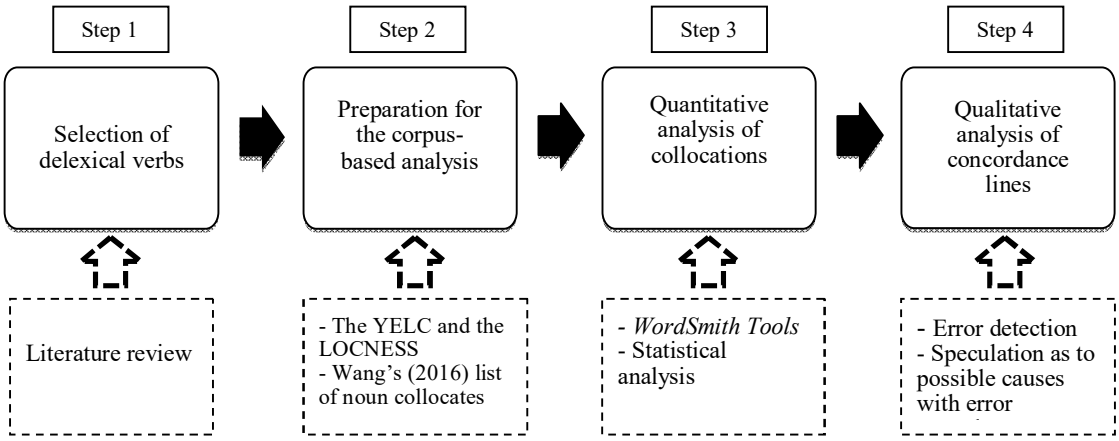

Figure 1. Data analysis procedure for delexical verb + noun collocations.

In Step 3, I carried out a quantitative analysis to answer RQ 1 and RQ 2. I used the corpus software program *WordSmith Tools 5.0* to obtain frequencies of the collocations. First, I have retrieved all the concordance lines that contain the four verbs, and compared the noun collocates of them with Wang's list. Since this list contains considerable number of noun collocates, the collocations that match with the list can be considered as delexical collocations. Then, I conducted statistical analyses to find significant differences between the YELC and the LOCNESS. To do so, I used an online corpus frequency test wizard: Statistic Inference: A Gentle Introduction for Linguists (SIGIL)[5]. This website provides an online tool for chi-square test (and G test, when needed) between two samples.

In Step 4, I examined the deviant collocations of delexical verb + noun constructions through a qualitative analysis. In order to show possible causes of the deviant collocations and thus answer RQ 3, I looked through only those concordance lines retrieved from the low level sub-corpus in the YELC (i.e., 359 concordance lines, see Table 5). I focused on the low level not only because of time constraints, but also because L2 learners at low proficiency level are more likely to produce deviant collocations. In order to determine deviant collocations, I used the BNCweb reference corpus as a norm. In corpus linguistics, the rate of co-occurrence between two or more words can be statistically measured, and the rate is used as a cut-off point to determine "collocation". A t-score of 2 or higher as a co-occurrence rate is normally regarded to be significant (Hunston, 2002). I checked t-scores for delexical collocations through the BNCweb, and found several collocational errors. In so doing, I was able to find that quite a few combinations of the four verbs with noun collocates were deviant.

Results and Discussion

Comparison of Frequency of Delexical Collocations Between the YELC and the LOCNESS

Overall frequency of the four high-frequency verbs

Before analyzing the frequencies of delexical collocations of the four high-frequency verbs, this study investigated whether L2 learners under- or overuse the given verbs in comparison with native speakers. As shown in Table 2, *WordSmith Tools* retrieved 11,846 occurrences in the YELC and 2,581 occurrences in the LOCNESS. In terms of the normalized frequency (occurrence per 1,000 words), Korean learners

[5] It can be accessed at http://sigil.collocations.de/wizard.html.

used the verbs more frequently than native speakers (10.91 vs. 7.96). The result of the chi-square test shows that the difference between the two groups is statistically significant (chi-square = 214.34715, $p < .001$), which means that Korean learners use the four high-frequency verbs significantly more than native speakers do.

TABLE 2
Frequency of the High-Frequency Verbs in the YELC and the LOCNESS

	The YELC	The LOCNESS
Occurrence of the four verbs	11,846	2,581
Overall token of the corpus	1,085,927	324,304
Occurrence per 1,000 words	10.91	7.96

It is difficult to speculate as to why, as found in the present study, L2 learners significantly overuse the high-frequency verbs. Moreover, the results from previous studies are not consistent, and are sometimes conflicting. In the analysis of the high-frequency verb *make*, Altenberg and Granger (2001) showed that Swedish learners significantly overused *make* whereas French learners of English underused it in comparison with native speakers. Interestingly, previous research has shown that Korean learners of English overuse the high-frequency verbs. According to Ha and Nam (2018), Korean university students significantly overused the four high-frequency verbs (*make*, *take*, *give*, and *get*) in comparison with native speakers, which is line with the present study. Similarly, Lee and Na (2015) reported that Korean learners overused *make*, and that this result was related to a causative use of the verb. One possible reason for these conflicting results could be the different effects of L1 transfer for L2 learners with different L1 backgrounds. However, the present study focuses on delexical collocations of the four high-frequency verbs, rather than on under- or overuse of the verbs. Note that the overall frequency of these verbs in Table 2 includes not only the occurrence in delexical collocations, but also other usages of the verbs.

Comparison of delexical verb + noun collocations between the YELC and the LOCNESS

As explained earlier, this study detected occurrences of delexical verb + noun collocations in both the YELC and the LOCNESS, based on Wang's (2016) list of noun collocates. Table 3 presents the results of this comparative analysis. First, the number of occurrence (raw frequency) of delexical collocations in the YELC is greater than that in the LOCNESS. This is not surprising because the YELC is much larger than the LOCNESS. However, the occurrence rates of the two groups (16.8% and 28.9% respectively) indicate that Korean learners produce fewer delexical collocations than native speakers. Accordingly, it can be argued that when Korean learners use the high-frequency verbs, they are more likely to use them with literal senses rather than in delexical constructions. Furthermore, the difference in occurrence rates between the two groups is statistically significant (G square = 184.30340, $p < .001$).

TABLE 3
Frequency of Delexical Verb + Noun Collocations in the YELC and the LOCNESS

	The YELC	The LOCNESS
Overall occurrence of the four verbs (a)	11,846	2,581
Number of occurrences of delexical verb + noun collocations (b)	1,995	746
Occurrence rate (%)[6]	16.8%	28.9%

[6] The formula for this measurement = (b) / (a) * 100. One reviewer has mentioned that it would be good to include the other kind of occurrence rates that can be calculated by using the total words of each corpus as a denominator, rather than the occurrence of the four words. This may show a different aspect of the results, which I do not think is necessary to be included. By the way, using the total words as a denominator, the difference between the YELC and the LOCNESS is also statistically significant (0.18% vs. 0.23%, $p < .001$).

Considering that, as reported in the previous section, Korean learners significantly overuse the four high-frequency verbs, it is interesting that they underuse delexical collocations. This result echoes the findings of previous studies. For example, Altenberg and Granger (2001) showed that Swedish and French learners of English underused delexical uses of the high-frequency verb *make*. Likewise, Gonzalez Alvarez and Doval Suarez (2010) revealed that Spanish learners of English underused free combinations and collocations of *take*, compared with native speakers. In contrast, Wang (2016) reported that Chinese and Swedish learners of English overused delexical collocations compared with native speakers, which is in conflict with the present study. However, the L2 learners in Wang's study appeared to use only a limited variety of delexical collocations, which suggests that L2 learners repeatedly use certain delexical collocations.

The effect of individual verbs

The findings presented above show that the frequency of delexical collocations made by Korean learners is significantly different from that of native speakers. However, if the occurrence of delexical collocations that involve one or two idiosyncratic high-frequency verbs is extraordinarily high or low, this could skew the findings of significant difference between the two groups. In order to check whether the findings might be misleading, this study examined the frequencies of delexical collocations of each individual verb, as shown in Table 4.

As can be seen from Table 4, for each individual verb the raw occurrence of delexical constructions in the YELC is much greater than that in the LOCNESS. However, in almost every case the occurrence rate of delexical collocations is lower in the YELC than in the LOCNESS, although there is a degree of difference between individual verbs. The one exception is the verb *get*, where the occurrence rate is slightly higher in the YELC than in the LOCNESS.

It is not immediately clear why the verb *get* is an outlier in these results. One possible reason may be related to the small size of the LOCNESS. That is, the raw occurrence of delexical collocation of *get* in the LOCNESS may be too low to be comparable (only 37 occurrences). However, whatever the reason might be, when taking the total difference between the two corpora into account (16.8% vs. 28.9%) it appears that the impact of this outlier is limited. Hence, it can be argued that there is no significant effect from any individual verb. Therefore, the finding that the frequency of delexical collocations made by Korean learners is significantly different from that of native speakers does not appear to be misleading, and the tendency of Korean learners to underuse delexical verb + noun collocations is consistent, despite the existence of an outlier.

TABLE 4
Comparison of Individual Delexical Verb + Noun Collocations

Individual verb	The YELC			The LOCNESS		
	Overall occurrence	Occurrence of delexical collocations	Occurrence rate (%)	Overall occurrence	Occurrence of delexical collocations	Occurrence rate (%)
make	5,582	502	9.0%	1,017	279	27.4%
take	1,766	677	38.3%	675	266	39.4%
give	2,043	582	28.5%	466	164	35.2%
get	2,455	234	9.5%	423	37	8.7%
Total	11,846	1,995	16.8%	2,581	746	28.9%

Frequency of Delexical Collocations According to English Proficiency Level

In order to show whether and/or how L2 learners develop a native-like command of delexical verb + noun collocations, this study investigated the frequencies of the collocations in each sub-corpus of the YELC: low, intermediate, and advanced proficiency levels (see Table 1).

Table 5 shows the occurrence rates of delexical collocations in the YELC according to proficiency level. The fact that the occurrence rate in the LOCNESS (28.9%) is significantly higher than that in the YELC (16.8%) (see Table 3) indicates that native speakers tend to use delexical collocations more frequently than L2 learners. Accordingly, it could be predicted that as L2 learners' proficiency in English improves, the occurrence rate of delexical collocations should increase accordingly. As shown in Table 5, the rate rises slightly from low level (15.8%) to intermediate level (17.1%). However, at advanced level it decreases a little to 16.1%, which is only 0.3% point higher than the rate at low level. A series of chi-square tests indicate that the differences between any pair of two proficiency groups is not statistically significant.

TABLE 5
Frequency of Delexical Collocations in the YELC According to English Proficiency Level

	Low	Intermediate	Advanced
Overall occurrence of the four high-frequency verbs	2,269	9,005	572
Occurrence of delexical collocations	359	1,544	92
Occurrence rate (%)	15.8%	17.1%	16.1%

These results suggest that it would be difficult for L2 learners to acquire a native-like command of delexical collocations in terms of frequency, even when they reach an advanced level of English proficiency. Hence, they offer valuable implications for language teachers and learners in the English classroom. For example, language teachers may be required to provide explicit instructions to teach delexical collocations to L2 learners. L2 learners should be exposed to various kinds of delexical collocations of the high-frequency verbs in order to raise their awareness.

Possible Causes of the Collocational Errors

The sentences in (1) are examples[7] in which Korean learners have chosen appropriate verbs for the given noun collocates. For example, in (1a) a learner appears to have known that the noun collocates "effort" and "mistake" adopt the high-frequency verb *make* to produce relevant delexical collocations. Likewise, examples (1b), (1c) and (1d) show the correct combinations of delexical verbs with the given noun collocates.

(1) a. Just with these penalties, students make an effort not to make a mistake.
b. We should take actions to avoid accidents.
c. In fact, I want to be a person who gives a smile to other people like Ji Min.
d. I want to travel Japan if I get a chance to go abroad.

However, as mentioned earlier, this study looked into collocational errors involving delexical verbs and noun collocates, as found in the low level sub-corpus of the YELC. As with previous research that has shown interlingual and intralingual influences, the collocational errors in the YELC can be divided into two different kinds according to their possible causes.

First, the present study found that Korean learners produced several collocational errors that appear to be literally translated from the L1 (Korean). The example sentences in (2) seem to be evidence of interlingual influence, where learners have chosen inappropriate verbs for given noun collocates. For example, in (2a) a Korean learner has chosen the verb *do* rather than the delexical verb *take* for an object "actions." This might be because the collocation *take action* is replaced by a literal translation of the Korean *행동하다,* i.e., **do action*. Similarly, the collocation **do a effort* in (2b) also seems to be derived

[7] Note that the examples are retrieved from writings made by Korean learners. Hence, they may contain spelling, grammatical and lexical errors as well as deviant delexical collocations.

from the Korean expression. The example (2c) is interesting in that the learner has produced a very odd collocation **eat a drug*, for the target expression *take a medicine*. This deviant expression also seems to stem from its corresponding expression (i.e., *약을 먹다*) in the Korean language. The examples (2d) and (2e) also appear to be the results of literal translation. In other words, these collocational errors can be attributed to non-congruence between L1 and L2 (Nesselhauf, 2005), and thus provide evidence of interlingual influence.

(2) a. *Because most people do bad actions on the Internet like bad replies that make people die. (<*take action*)
b. *so although it is very hard, smokers are doing a effort. (<*make an effort*)
c. *Actually many students speak headache, stomachache, dry eyes, and many sick. So, they usually eating a drug. (<*take a medicine*)
d. *She usually said jokes that all my class student was made laugh. (<*make a joke*)
e. *Of course, they can have rights that catching a very emergency call, but when it happens, they should stop driving on the street side first and catch a call. (<*take a call*)

On the other hand, this study also found quite a few collocational errors that seem to be affected by intralingual influence. The deviant collocations from (3) to (6) have been selected as examples for each of the four delexical verbs (*take*, *make*, *have*, and *get*, respectively). They show that Korean learners appear to be confused when choosing appropriate delexical verbs, probably because these verbs are used with little or no meaning (Chi et al., 1994).

For example, the sentences in (3) include collocational errors in which the target verb is *take*. Korean learners made deviant collocations like **have an action* as in (3a) or **have a photo* as in (3d). According to the BNCweb, the t-scores of these collocations are under 2, hence these are considered to be collocational errors.

(3) a. *I want to have more specific thing but actually I don't have any action. (<*take action*)
b. *Consequently it can make other more serious similar action. (<*take action*)
c. *But recently, as they feel the danger and seriousness, they make many cares and regulations. (<*take care*)
d. *And I can see my favorite baseball player, if I'm lucky, I have a photo with him. (<*take a photo*)
e. *Finally I wonder how much pride the english have. (<*take pride*)

The sentences in (4) are examples of collocational errors in which the target verb is the delexical verb *make*. Learners have made deviant delexical collocations like **take a mistake* as in (4a) or **have a mistake* as in (4b), rather than the target form *make a mistake*. They have also created deviant expressions like **have a trip* (for *make a trip*), **get an adjustment* (for *make an adjustment*), and **take a change* (for *make a change*) as in (4c), (4d) and (4e) respectively. Since the delexical verb *make* in these collocations does not have a core sense of "produce," but retains little literal meaning, Korean learners would not be able to recognize that the given noun collocates should adopt this specific delexical verb. Rather, they are as likely to use other delexical verbs, like *take*, *have*, or *get*. It seems that these errors are not affected by interlingual influence, because they are not caused by literal translation from Korean. If L1 transfer were to play a role in causing errors with the delexical collocation *make a mistake*, Korean learners would produce **do a mistake* (i.e., *실수하다*) rather than **take a mistake* as in (4a).

(4) a. *If Drivers use cellular phones while driving, drivers take a mistake on driving. (<*make a mistake*)
b. *I think that when they have big mistakes, and they don't apologize, physical punishment is needed just a little bit. (<*make a mistake*)

c. *And sometimes, our family go out for a dinner or have a trip for a short time. (<*make a trip* or *take a trip*)
d. *If a person get an adjustment, he could be a harmful man to society. (<*make an adjustment*)
e. *For them, our government should take a change of law. (<*make a change*)

Similarly, the delexical collocations presented in (5) and (6) show evidence of intralingual influence rather than interlingual influence. The sentences in (5) are examples of collocational errors involved with the target verb *have*, whereas the sentences in (6) are related to the verb *get*.

(5) a. *By that day, I get confidence for being in front of many audience. (<*have confidence*)
b. *When we make some ideas for the cure of disease, we need to take experiment this theory. (<*have an idea*)
c. *So, I could get a hope again and regained my confidence. (<*have a hope*)
d. *One or two hours later, I make a meet with my friends. (<*have a meeting*)

(6) a. *All spicies are selfish for making their benefit of life, and that way, they survive. (<*get a benefit*)
b. *… they comment on their real name, but only open name, they only have blames but can't keep them stop. (<*get blame* or *take blame*)
c. *I has taken his help in the very variety situation too. (<*get help*)
d. *They take a shock who read words are harmful about themselves so are taken care of by counselors to protect their suicides. (<*get a shock* or *have a shock*)
e. *… what is the difference of taking a deep sleep which driving, or taking a short nap while driving. (<*get a sleep* or *have a sleep*)

In summary, according to the qualitative analysis, Korean learners of English make many errors of delexical verb + noun collocations, which can be divided into two kinds according to their possible causes. Some errors appear to be affected by L1 transfer (interlingual influence), through literal translation from the learners' L1. The other errors seem to derive from learners' incorrect choices of delexical verbs, i.e. the overgeneralization of particular delexical verbs (intralingual influence). These findings echo the results of previous studies that have shown both interlingual and intralingual influences on collocational errors made by L2 learners. The present study is noteworthy in that it is probably the very first research to speculate as to the possible causes of deviant delexical collocations made by Korean learners of English.

Conclusion

This study aimed to show how delexical verb + noun collocations made by Korean learners of English differ from those made by native speakers. It adopted a corpus-based analysis to investigate whether Korean learners under- or overuse delexical collocations involving four high-frequency verbs (*make*, *take*, *give*, and *get*), and the possible causes affecting the collocational errors.

The findings gained through quantitative analysis of and comparison between the YELC and the LOCNESS show that Korean learners significantly underuse delexical collocations although they significantly overuse the high-frequency verbs in general, which is in line with previous studies (Altenberg & Granger, 2001; Gonzalez Alvarez & Doval Suarez, 2010). This suggests that Korean learners might have difficulty in producing delexical collocations. Furthermore, it seems that Korean learners find it difficult to acquire a native-like command of delexical collocations. This is supported by the fact that the occurrence rates across three levels of English proficiency are not significantly different. Even Korean learners at an advanced level significantly underuse delexical collocations, compared with native speakers.

The qualitative analysis of the errors of delexical collocations made by Korean learners suggests that deviant delexical collocations might be due to either interlingual or intralingual influence. In line with the findings of previous studies, this study has shown that Korean learners sometimes literally translate Korean expressions into deviant delexical collocations in English (interlingual influence). They also select inappropriate delexical verbs, which results in collocational errors (intralingual influence).

The findings of this study offer valuable implications for second language classrooms. For example, as Korean learners appear to underuse delexical collocations and have difficulty in acquiring a native-like command of this feature, language teachers may have to provide implicit and/or explicit instructions to improve learners' understanding of delexical verb + noun collocations. These kinds of instructions should be given not only to learners with low English proficiency, but also to those at intermediate and advanced levels. Language teachers should also be aware that learners could make collocational errors affected by both interlingual and intralingual influences, and may require relevant instructions.

One of the limitations of the present study is related to the size of the learner corpus. Although the YELC consists of approximately 1 million words, which might be considered quite large, the results would be more robust if a larger learner corpus was analyzed. However, it should also be noted that the YELC is the largest learner corpus currently available in the Korean context. Another limitation is that the present study investigated only the low level sub-corpus to identify possible causes of delexical collocations because of time constraints. However, considering the occurrence rates of delexical collocations according to proficiency levels are quite similar (see Table 5), it would also be worth examining intermediate and advanced level sub-corpora.

Finally, pertaining to the strengths and the limitations of the present study, it can be suggested that future research can investigate data from higher proficiency level learners to delve into the possible causes of deviant delexical collocations. It is also suggested that similar types of research should be conducted with naturally occurring data made by secondary or primary school students, in order to see the developmental patterns through a longitudinal or quasi-longitudinal study. This study hopes to contribute to the understanding of Korean learners' acquisition of delexical collocations.

The Author

Wonkyung Choi is a PhD candidate in the Department of English Language and Linguistics of the University of Birmingham in UK. His current research interests cover corpus linguistics, learner corpus research, English language teaching (ELT), and second language acquisition.

Department of English Language and Linguistics
College of Arts and Law
The University of Birmingham
Edgbaston, Birmingham, B15 2TT, UK
Mobile: + 82-2-2243-6108
Email: yolsimi77@gmail.com

References

Allan, Q. (1998). Delexical verbs and degrees of desemanticization. *Word, 49*(1), 1-17.

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.

Babanoglu, M. P. (2014). A corpus-based study on the use of MAKE by Turkish EFL learners. *International Journal of Education & Literacy Studies, 2*(2), 43-47.

Chi, A. M. L., Wong, K., & Wong, M. (1994). Collocational problems amongst ESL learners: A corpus study. In L. Flowerdew & A. Tong (Eds.), *Entering text* (pp. 157-165). Hong Kong: Language Centre, The Hong Kong University of Science and Technology.

Choi, S. (2012). *A corpus-based analysis of the light verb collocation in English textbooks published by EBS* (Unpublished master's thesis). Sogang University, Korea.

Gonzalez Alvarez, E. M., & Doval Suarez, S. M. (2010). TAKE + NOUN sequences in native and learner written data. *Revista Electronica de Linguistica Aplicada, 10*, 55-68.

Gouverneur, C. (2008). The phraseological patterns of high-frequency verbs in advanced English for general purposes: A corpus-driven approach to EFL textbook analysis. In F. Meunier & S. Granger (Eds.), *Phraseology in foreign language learning and teaching* (pp. 223-243). Amsterdam: John Benjamins Publishing Company.

Granger, S., Dagneaux, E., Meunier, F., & Paquot, M. (2009). *The international corpus of learner English* (Version 2). Handbook and CD-Rom.

Ha, M-H., & Nam, J-M. (2018). Learner corpus-based study on English light verbs. *Journal of Language Science, 25*(1), 269-287.

Hill, J. (2000). Revising priorities: From grammatical failure to collocational success. In M. Lewis (Ed.), *Teaching collocation: Further development in the lexical approach* (pp. 47-69). Boston, MA: Heinle.

Hong, A. L., Rahim, H. A., Hua, T. K., & Salehuddin, K. (2011). Collocations in Malaysian English learners' writing: A corpus-based error analysis. *The Southeast Asian Journal of English Language Studies, 17*, 31-44.

Howarth, P. (1996). *Phraseology in English Academic Writing: Some implications for language Learning and dictionary making*. Tübingen: Niemeyer.

Hunston, S. (2002). *Corpora in applied linguistics*. Cambridge: Cambridge University Press.

Jang, E. Y. (2008). *An analysis of collocation in the elementary ELT reading textbooks: With a focus on delexical verbs*. MA thesis. Yonsei University.

Jespersen, O. (1965). *A modern English grammar on historical principles: Part VI, Morphology*. London: George Allen and Unwin Ltd.

Jukneviciene, R. (2008). Collocations with high-frequency verbs in learner English: Lithuanian learners vs native speakers. *KALBOTYRA, 59*(3), 119-127.

Kang, M-K. (2014). A corpus-based analysis of Korean high school English textbooks based on the semantic distribution of high-frequency verb collocation. *English Language Teaching, 26*(1), 113-130.

Kim, M-H. (2002). Uses of *Make* in Korean EFL learner writing: A corpus-based study. *English Teaching, 57*(4), 297-314.

Kittigosin, R., & Phoocharoensil, S. (2015). Investigation into learning strategies and delexical verb use by Thai EFL learners. *3L: The Southeast Asian Journal of English Language Studies, 21*(2), 63-72.

Kwon, H-Y. (2012). *An analysis of 'get, have, take' collocations in high school 10 textbooks corpus-based study* (Unpublished master's thesis). Kyung Hee University, Korea.

Laufer, B., & Waldman, T. (2011). Verb-noun collocations in second language writing: a corpus analysis of learners' English. *Language Learning, 61*(2), 647-672.

Lee, S., & Na, Y-H. (2015). The use of the verb *Make* in the corpora of American NS students and Korean EFL learners. *English 21, 28*(1), 401-426.

Lennon, P. (1996). Getting 'easy' verbs wrong at the advanced level. *IRAL, 34*(1), 23-36.

Lewis, M. (1993). *The lexical approach: The state of ELT and a way forward*. Hove: Language Teaching Publications.

Liu, C-P. (2010). EFL learners' use of lexical collocations involving the light verbs *Have*, *Make*, and *Take*. *Hwa Kang English Journal, 16*, 91-116.

Nation, I. S. P. (2001). *Learning vocabulary in another language*. Cambridge: Cambridge University Press.

Nesselhauf, N. (2003). The use of collocations by advanced learners of English and some implications for teaching. *Applied Linguistics, 24*(2), 223-242.

Nesselhauf, N. (2005). *Collocations in a learner corpus*. Amsterdam: John Benjamins Publishing Company.

Rhee, S-C., & Jung, C. K. (2014). Compilation of the Yonsei English Learner Corpus (YELC) 2011 and its use for understanding current usage of English by Korean pre-university students. *The Journal of the Korean Contents Association, 14*(1), 1019-1029.

Seo, A. R. (2011). *A corpus-based analysis of delexical verb collocations in high school English textbooks for the first grade* (Unpublished master's thesis). Kongju National University.

Shin, E. J. (2012). *A corpus-based analysis of the delexical verb HAVE in middle & high school English textbooks* (Unpublished master's thesis). Kyung-Hee University.

Sinclair, J. M. (1990). *Collins Cobuild English grammar*. London: Harper Collins.

Vrbinc, M. (2004). Delexical structures contrastively: A common trap for non-native speakers. *English Language Overseas Perspectives and Enquiries, 1*(1-2), 61-78.

Walker, C. (2008). *A corpus-based study of the linguistic features and processes which influence the way collocations are formed* (Unpublished doctoral dissertation). University of Birmingham.

Wang, Y. (2016). *The idiom principle and L1 influence: A contrastive learner-corpus study of delexical verb + noun collocations.* Amsterdam/Philadelphia: John Benjamins Publishing Company.

Yan, Q. (2006). A corpus-based analysis of the verb "DO" used by Chinese learners of English. *CELEA Journal, 29*(6), 37-41.

Zheng, L., & Xiao, R. Z. (2015). A corpus-based study of collocation in Chinese EFL learners' oral production. *Corpus Linguistics Research, 1*, 83-108.