THE JOURNAL OF ASIA TEFL Vol.10, No. 2, pp. 63-89, Summer 2013

A Corpus-based Study of the Use of Phrasal Verbs in Korean EFL Students' Writing[1]

Mi-Lim Ryoo
Korea Maritime University, South Korea

This study sheds light on one of the most creative, thus challenging, grammatical classes of the English language, namely phrasal verbs (PV) in a corpus of Korean EFL students' writing. This aims to contribute a further facet to the general understanding of Korean learners of English. Taking the perspective provided by a usage-based approach to language, a corpus was compiled from 257 Korean students' essays from essay contests. The most frequent verbs and adverbial particles in PV constructions and PVs in the corpus were identified and then compared to the actual use of native speakers of English in the British National Corpus. It was found that the top 4 most frequent verbs (e.g. GO, COME) and adverbial particles (e.g. up, out) in both corpora were almost identical, and more than half of the top 20 verbs overlapped. The most frequent PV combinations in the BNC also often occurred in the student corpus, while less frequent combinations in the former rarely appeared in the latter. These findings can provide evidence indicating that Korean EFL students lack the formulaic competence of PVs. The implications of the findings for English language teaching and learning are also discussed.

Keywords: phrasal verbs, lexical verbs, adverbial particles, Korean EFL students' writing, usage-based approach

---

[1] This work was supported by the Korea Maritime University Grant.

INTRODUCTION

Over the past four decades, a growing number of corpus linguistic research has shown that whether they are idioms (e.g., *a piece of cake*), collocations (e.g., *blonde hair*), or prefabs (e.g., *the thing is*), fixed or semi-fixed phraseological units[2] (or multi-word expressions) are pervasive in English as well as many other languages (e.g., Altenberg, 1998; Biber, Johansson, Leech, Conrad, & Finegan, 1999; Sinclair, 1991, 1996). Altenberg (1998) assumes that the lexicon of adult native speakers may consist of up to 80 percent of such phraseological units. Biber et al. (1999) identify a large number of lexical bundles (in their terms) in both spoken and written texts. Innumerous corpus linguistic research has contributed to a high degree to identifying phraseological items, patterns, and constructions. Thus, as far as fluency and comprehension are concerned, acquiring phraseological units is important for both native speakers and non-native speakers alike (Nesselhauf, 2003, 2005).

One of the most important areas for progress in the discussion of phraseological units is phrasal verbs[3] which constitute one of the most distinctive and creative features of the English language (Gardner & Davies, 2007). Because phrasal verbs (thereafter PVs) consist of one open-class item (the verb) and one closed-class item (the particle), they are at the interface of lexis and grammar "that has important ramifications for second language acquisition" (Gardner & Davies, 2007, p. 340; Gass & Selinker, 1992; Howarth, 1998). PVs are typical of spoken and informal English, but also widely used in written and formal English (Fletcher, 2005; McCarthy & O'Dell, 2004). When it comes down to non-native learners of English, particularly learners with non-Germanic first languages (L1s), PVs are notoriously difficult to acquire, especially due to the fact that they tend to permeate and be

[2] Cowie (1998) states that "[p]hraseology is a field bedeviled by the proliferation of terms and by conflicting uses of the same term" (p. 210). Wray (2002) lists more than 50 terms used for phraseological units. In this paper, the terms 'multi-word expression' or 'phraseological unit' will be used as these are fairly neutral terms that do not refer to any particular theoretical framework.

[3] Phrasal verbs are not always considered phraseological units in the literature (See Waibel , 2007, for details).

highly productive in English (Celce-Murcia & Larsen Freeman, 1999; Darwin & Gray, 1999; Moon, 1998). Sinclair (1996) even refers to them as "the scourge of the learner" since they present so many inherent difficulties, such as "idiomaticity or polysemy" (p. 78). PVs therefore present a worthy field of study as far as non-native learners of English are concerned.

In spite of the importance of PVs in gaining native like fluency, little research has been done on the use of PVs by Korean learners of English. Only in recent years, a few researchers have shown interest in PVs in the Korean context (An, 2012; Sung, 2012), but to our knowledge, no research has dealt with the actual use of PVs by Korean EFL students. In this regard, the present paper investigates the use of PVs by Korean EFL students via an analysis of a student corpus, asking more practical questions about what kinds of PVs Korean EFL students use, how frequently they use them, and what pedagogical implications it conveys. More precisely, this paper attempts to give evidence indicating what are the most frequent PVs, lexical verbs in PV constructions, and adverbial particles in Korean EFL students' writing. It also compares and contrasts the use of PVs between Korean students and native speakers of English.

The present study takes the perspective provided by a usage-based approach to language which "emphasize[s] the notion that actual language use is a primary shaper of linguistic form and the foundation for language learning" (Tyler, 2010, p. 270). That is to say, the study is concerned with learner performance rather than with theories on the semantic and syntactic aspects of phrasal verbs.

With its detailed analyses of phrasal verb use in Korean students' writing, the present study aims at contributing a further facet to the general understanding of Korean learners of English. In other words, this paper is intended to provide some empirical data for an understanding of Korean EFL students in the context of English teaching and learning; where they stand at the moment, and what they need to be more fluent English users, in terms of PV use. As for a researcher and teacher, it is believed that seeking for useful methods for effective teaching would be the upmost important task, but knowing her students' status quo is equally important for effective teaching. As the present research is based on the extraction of all PVs

rather than only those from a predefined list, a great advantage of this study is to provide all instances of PVs that Korean EFL students use.

LITERATURE REVIEW

In recent years, phraseology, "the study of the structure, meaning, and use of word-combinations" (Cowie, 1994, p. 3168), which had been long regarded as a peripheral issue[4] seems to take central stage in applied linguistics as well as second language acquisition (SLA) research (Granger & Meunier, 2008; Gries, 2008). The current status of phraseology was originated with John Sinclair's pioneering lexicographic work (Granger & Paquot, 2008; Römer, 2009). Sinclair's (1991, 1996) idiom principle (or "phraseological tendency") in particular describes how words do not occur in isolation but "go together and make meanings by their combinations" (2004, p. 29), and it has generated a recent frequency-based phraseological approach to language (Granger & Paquot, 2008). Sinclair's emphasis on phrase is further manifested in his conference presentation paper in 2008, entitled 'The phrase, the whole phrase, and nothing but the phrase.' Here he contends that "[w]e have to concede that the normal primary carrier of meaning is the phrase not the word" (p. 409). Perhaps Sinclair's finding of extended lexical units was possibly thanks to the advances of computer technology that allowed him to handle large electronic data of actual language, so-called corpus, plural corpora, but it was Sinclair himself who made the observation that "required the stroke of genius" (Stubbs, 2009, p. 22).

As Römer (2009) states, one significant finding of corpus linguistic research is that language is highly patterned. In other words, electronic corpora have enabled researchers to identify and classify phraseological units, including PVs and otherwise elusive structures that are pervasive in the English language (Hunston, 2002; Moon, 1998; Stubbs, 2001). Among various phraseological units, such as

[4] Kennedy (2008) provides a thorough explanation why and how phraseology had been neglected in linguistics. See also Ellis (2008).

idioms, prefabs, and collocations, the PV classified as the category of collocation is particularly important because not only does it constitute a major grammatical class in native speakers' discourse, but also it lies at the interface of lexis and grammar.

Although phrasal verbs are "a vigorous part of English," they had been long neglected by scholars (McArthur, 1989), and only in the 20th century have they received the comprehensive description that they deserved (Perdek, 2010). In spite of their relatively recent establishment as a grammatical category, however, there is an extensive literature on PVs' syntactic (e.g., Dehé, 2002; Fraser, 1976; Sroka, 1972), semantic (e.g., Bolinger, 1971), and pragmatic functions (O'Dowd, 1998). Likewise, "the extreme heterogeneousness" of their forms and "absolute unpredictability" in both formation and meaning (Frank, 1989, p. 137) have made PVs "special enough to warrant production of specialized dictionaries" (Perdek, 2010, p. 1391). According to Perdek, the first dictionary of phrasal verbs was compiled by George Mayer and published in 1975.

PVs are a type of multi-word verb. All multi-word verbs constitute a lexical and syntactic unit, functioning as a single lexical verb. They consist of a verb proper and one or two additional elements, called particles. Quirk, Greenbaum, Leech, and Svartvik (1985) divide multi-word verbs into phrasal verbs, prepositional verbs, and phrasal-prepositional verbs. The subdivisions of multi-word verbs constituents are associated with the nature of the particle. As Perdek notes, whether the particle is a preposition or an adverb or both is an ongoing discussion. The general consensus, however, is that in phrasal verbs, the particle is an adverb (e.g., *bring back*), in prepositional verbs a preposition (e.g., *listen to*), and in phrasal-prepositional verbs an adverb and a preposition (e.g., *put up with*). It is not our task in this study to classify PVs in any way; rather, our aim is solely pedagogical, simply to identify what PVs are used and to what extent they are used. Thus, the the term particle will be used to avoid unnecessary confusion.

The phrasal verb is "a fuzzy grammatical category" that all dictionaries were inconsistent in terms of providing grammatical objects in its definition, as a result, locating it as an independent category in some dictionaries or as a subcategory of either the verb or the preposition in others (Gardner & Davies, 2007, p. 341; Perdek, 2010). The very name for phrasal verb is also controversial. Among various names,

such as "separable verb" (Francis, 1958), "verb-particle collocations" (Sroka, 1972), "verb-particle combinations" (Fraser, 1976), and "two-word verb" (Meyer, 1975), "phrasal verb" "appears (…) to be the winning term" (McArthur, 1989, p. 38). Hence, phrasal verb will be used in this paper since it also predominates in current student grammar books and teaching materials.

For purposes of the current study, we will rely on a more functional and objective definition of PV, that is, all two-part verbs consisting of a lexical verb (LV) proper (tagged as VV (VVI, VVB, VVD, and so on) in the corpus) followed by an adverbial particle (tagged as AVP) that is either contiguous (adjacent, such as *go back*) to that verb or noncontiguous (i.e., separated by one or more intervening words, such as *give* it *up*, *put* my name *on*). Furthermore, this study is concerned with actual learner performance rather than with a theoretical discussion of phrasal verbs.

Most SLA research on the use of PVs by learners of English thus far seems to fall largely in one of two methodological approaches, either investigating overuse or underuse by comparing the quantity of PVs in a native English speakers corpus and a L2 learners (or learners' textbooks) corpus (Durrant & Schmitt 2009; Sung, 2012) or examining learners' preference between a single lexical verb and a PV equivalent (Dagut & Laufer, 1985; Hulstijn & Marchena, 1989; Laufer & Eliasson, 1993; Liao & Fukuya, 2004). Durrant and Schmitt (2009) examined the use made by learners of collocations with various levels of frequency in the BNC. They found that adult non-native learners do make extensive use of collocations that are frequent in English, but fail to use less frequent, strongly associated collocations, items that are probably highly salient for native speakers. In other words, non-native learners also use phraseological units in a broad sense that are frequent in English, but do not use less frequent ones, which indicates that learners need substantial exposure to the target language in order to build up knowledge of a large number of phraseological units.

Sung (2012) compiled a corpus of 12 high school English textbooks in Korea to identify the types of PVs and their frequencies introduced in those textbooks. He found that a small subset of lexical verbs (e.g., *find, put, give, go, come*) combines with the most frequent particles (e.g., *up, on, out, in*), accounting of more than half

the phrasal verb occurrences. The most frequent top 10 PVs he identified are *find out* (18 times), *put on* (17), *give up* (11), *go on* (9), *come over* (8), *get on* (8), *look up* (8), *pick up* (8), *grow up* (8), and *build up* (7).

With respect to the avoidance issue, Dagut and Laufer's (1985) and Hulstijn and Marchena's (1989) study investigated the use of the English PV construction by Hebrew and Dutch learners respectively, and both concluded that L2 learners preferably used one-word verbs over multi-word verbs when PVs would be the typical choice by native speakers. Despite the same results, these two studies have a crucial difference in that Dagut and Laufer's study examined Hebrew learners whose L1 lacks the PV construction, while Hulstijn and Marchena's study had Dutch learners whose L1 also has the PV construction like English. Dagut and Laufer ascribe Hebrew learners' avoidance of PVs to the structural difference between L1 and L2. Hulstijn and Marchena, however, argue that even learners whose native language actually contains PVs (in this case, Dutch) often avoid using such forms when communicating in English, finding the reasons in L1-L2 similarity and semantic complexity of L2. In much the similar vein, Wray (2002) claims that the adult second language learner primarily notices and remembers not phraseological units, but individual words.

Grounding the research on the findings of Dagut and Laufer's (1985) and Hulstijn and Marchena's (1989) study, Liao and Fukuya (2004) examined the avoidance of English PVs with disparate proficiency levels (the advanced and intermediate) of Chinese students whose mother tongue lacks the PV construction. They found that there were some differences in the use of PVs between disparate proficiency levels; the intermediate (45% of the time) used considerably fewer PVs than did the advanced (75% of the time), preferring one-word verbs. Yet, the use of PVs by the advanced was still lower than that of native speakers. As such, they conclude that both the intermediate and advanced tend to avoid using PVs.

There are other SLA researchers who also have focused on the avoidance issue, but claim that their approach to the issue is distinctive in that they see the avoidance as the tendency to use on average fewer PVs than native speakers (Alejo, 2010, 2012; Cobb, 2003), rather than presuppose L1 knowledge or a strategic behavior on the part of the learner (i.e., underuse). Alejo (2012) contends that L2 learners rely

on a small group of PVs that is very frequent in their input, while showing difficulty with the ones in the low frequency band. According to Alejo, underuse of PVs is associated with "the lack of formulaic competence by L2 learners" (p. 39). In a similar vein, Goldberg (2009) claims that non-native learners seem to rely more on memorized chunks than on a productive use of this construction.

Gardner and Davies (2007) whose study motivated this very study, provide a distinct approach to PVs from the perspective of teaching and learning. The purpose of their study was to establish a basis for contents (what to teach) for the teaching of PVs rather than methodology (how to teach). They analyzed a large representative corpus of English - British National Corpus: World Edition (BNC) - to see frequencies of actual occurrence of verb-plus-particle combinations. They found that PVs constitute a major grammatical class. More precisely, they contend that "learners will encounter, on average, one in every 150 words of English they are exposed to" (p. 347). Hoping to be used in English language teaching, materials development, and testing, they provide lists of the most frequent particles and verb-plus-particle combinations in the BNC.

METHOD

Data

Data for this study was drawn from 257 essays written by the students in English essay contests held by a Korean university from 2009 to 2012. The contestants were undergraduate students from various disciplines and university years. Due to the nature of the contest which any member of the student body could participate, the skill level of the contestants was unknown, which, some may argue, can be a critical variable in this type of the study. A common sense assumption is that the majority of the contestants would be either intermediate or advanced in English composition because any student who has the courage to participate in such contest must have competence to some degree in their writing skills.

The contests were carried out in a computer laboratory for two hours including the instruction time. Each year a different essay topic, but all narrative writings, was given to the contestants on the site. The students worked on their essay using a word processing program on a computer. Except for the computer, they could not use any devices or aids. The computers they use had no Internet connection.

The chief reason for using contest essays for this study was due to the fact that they revealed the contestants' actual use of the English language, thus eliminating any outside help that one could easily get if s/he wanted. In other words, contest essays were regarded as more natural use of the target language.

With 257 essays, we were able to compile an 111,542 running word English essay contest (EEC) corpus. The average length of each essay was 434 words. Although the EEC corpus was not a large database compared to some corpora that were compiled for general purposes of language research, such as the Corpus of Contemporary American English (425 million words as of March, 2011), COBUILD corpus (525 million words as of 2005), British National Corpus (100 million words), it can be regarded as a "custom-made" corpus that provided specific information that this study attempted to obtain. Since the essays were already in electronic format, they were saved as plain text files that "will offer the maximum flexibility of use with different software suites" (O'Keeffee, McCarthy, & Carter, 2007, p. 8).

Data Analysis

Taking a usage-based approach that does not postulate an innate grammatical system (i.e., Universal Grammar) to explain the outcome of successful language acquisition (Zyzik, 2009), the present study focuses on the students' performance with the PV construction in the EEC corpus. We first embedded the part-of-speech (POS) annotation in the text using an annotation tool which automatically assigns POS tags to lexical units, called the constituent likelihood automatic word-tagging system (CLAWS) developed at Lancaster University. This tagger is known to have an accuracy rate of 97% on general written English, leaving only a 3% error rate (3 in 100 words).

Next, we gathered every instance from the data that were identified as the PV construction which a LV is followed by an AVP. The AVP could immediately follow the verb (LV + AVP) or after one or more intervening words as in the forms of LV + X + AVP (e.g., bring *it* back), LV + X + X + AVP (e.g., bring *the book* back), and so on. We confined our attention to two intervening words because 1) our subsequent analyses of the gathered data revealed almost no instance of phrasal verbs with longer separations (mostly immediately adjacent), and 2) the previous research shows that the PV construction that goes beyond more than two intervening words presents many false PVs (Gardner & Davies, 2007).

We then sorted out all inflectional forms of the same verb (lemma) from all the LVs in PV constructions and counted them together, separating them from different lemmas. For example, go, goes, going, went, and gone were grouped under the lemma GO and counted altogether. Thereafter, we use uppercase letters to refer to lemmas that embrace all the inflections of a verb (see Gardner & Davies, 2007; Stubbs, 2001).

We also used one of the commercially available concordance tools, WordSmith 5 (Scott, 2008) for concordance searches. This search produced concordance lines of each AVP as well as each LV in PV constructions for further analyses of the node words, such as classifying the grammatical function of right and left words of the node word. Since we were fully aware of methodological issues that have been frequently discussed in the literature – one of the issues pertains to automatic analyses (using by tagging systems) of potentially ambiguous phrases or structures in natural language use, we manually analyzed one by one, by paying a closer attention to every single occurrence of PV constructions in context.

All the results obtained by analyzing the EEC corpus were compared to BNC data reported by Gardner and Davies (2007) in order to find out the differences, if any, between the non-native and native speaker in the frequency of verb-plus-particle combinations. Thus, for the purpose of this study, Gardner and Davies's study was modified and utilized in the course of data analyses.

RESULTS AND DISCUSSION

Types and Frequency of Adverbial Particles (AVPs) in EEC

Particles followed by a lexical verb were tagged as AVP in the EEC corpus. Tagged AVPs show 12 different types in the EEC corpus. Table 1 below presents those 12 AVPs, their total frequency, occurrences as the AVP among other potential grammatical categories, such as the preposition (as in 'display *on* the table') and the noun (as in '*back* of the body'), and the percentage of each AVP, along with the data of AVPs from the BNC:

TABLE 1
Frequency of AVPs in EEC and BNC[5]

	EEC				BNC		
Form	total #	# as AVP	% as AVP	Form	total #	# as AVP	% as AVP
up	183	168	91.8	up	180,792	158064	87.4
out	99	84	84.8	out	149,727	145706	97.3
back	63	54	85.7	back	97,154	75233	77.4
down	33	32	97	down	91,832	72709	79.2
in	2,080	23	1.1	on	705,790	54956	7.8
around	68	19	27.9	off	67,479	37751	55.9
on	46	16	3.5	in	1,845,0770	34411	1.9
by	300	13	4.3	over	128,304	32526	25.4
about	440	10	2.3	about	190,615	12587	6.6
off	21	8	38.1	round	30,821	10895	35.3
along	25	7	28	around	43,391	10384	23.9

[5] BNC data here are modified from Gardner and Davies (2007).

through	83	6	7.2	through	81,184	5797	7.1
				along	18,555	4925	26.5
				by	504,969	371	0.1
				under	60,049	313	0.5
				across	24,053	13	0.1
total	3855	440	22.6*	total	4,219,792	656641	15.6*

*: Average of column

The comparison between AVPs identified in the EEC corpus and in the BNC reveals an apparent difference in that the latter has four more AVPs, i.e., *over, round, under*, and *across*. In other words, the Korean learners of English in this study never used these four AVPs. Unlike *under* (313 times) and *across* (13) that rarely occur as the AVP in the BNC, thus perhaps could possibly be disregarded, *over* and *round* occur much frequently marking 32,526 times (25.4%) and 10,895 times (35.3%) respectively. From the perspective of teaching and learning, it can be said that *over* and *round* functioning as the AVP in PV constructions should be taught to Korean learners of English either explicitly or implicitly, or both. Other than this, there seems to be not much difference between Korean learners and native speakers in terms of the frequency rate of AVPs. In fact, the top four AVPs in both corpora (i.e., *up, out, back*, and *down)* are identical. Thus, it can be said that the high productivity of these four particles in English PV constructions is acquired by Korean EFL students.

By (300), *on* (460), *about* (440) and *in* (2,080) are the most frequent particles in the EEC corpus, but they were used mostly as prepositions, showing only 4.3%, 3.5%, 2.3% and 1.1% of the time respectively used as a part of PV constructions, namely AVPs. These results have an implication for English teachers that Korean learners more often recognize *down, up, back* and *out* as AVPs and *by, on, about*, and *in* as prepositions. The analyses of the BNC, however, suggest that *on* and *about* occur as AVPs 7.8% and 6.6% of the time respectively, while *in* 1.9% of the time and *by* 0.1% of the time. This means that native speakers are most likely to recognize *in* and *by* as prepositions rather than AVPs in PV constructions, whereas

on and *about* can be still used as AVPs 8 and 7 in every 100 uses, which cannot be neglected. As such, Korean learners of English should be informed that the particles *on* and *about* can also be used in such PV constructions as *go on* (ranked as number 1 in the list of top 100 PV lemmas in the BNC), *come on* (number 14), *bring about* (number 44), and *come about* (number 83). The actual use of verb-plus-particle combinations in the EEC corpus will be discussed in the later section.

Frequency of Lexical Verbs (LVs) in EEC

Lexical verbs in the EEC corpus were tagged as one of five forms: VVB (base form of LV, except the infinitive; e.g., take, live), VVD (past tense form of LV; e.g., took, lived), VVG (-ing form of LV; e.g., taking, living), VVI (infinitive of LV), VVN (past participle form of LV; e.g., taken, lived), and VVZ (-s form of LV; e.g., takes, lives). Table 2 shows the frequency of all LV tokens, the frequency they occurred in PV constructions, and total occurrences of LV lemmas, along with the BNC data for comparison:

TABLE 2
Frequency of LVs in EEC and BNC[6]

	EEC			BNC		
	total #	# in PVs	% of PVs	total #	# in PVs	% of PVs
LV tokens	13,768	397*	2.9	10,404,107	518,923*	5.0
LV lemmas	361	90	24.9	19,682	1,572	8.0

*: a total number of the constructions of V +AVP, V + X + AVP, and V + X + X + AVP.
Lemma: all inflectional forms of a verb (e.g., go, goes, going, went, gone) considered to be the same verb (e.g., GO).

Compared to the percentage of PVs (5.0%) in the total LV tokens in the BNC, the EEC corpus presents a lower percentage (2.9%). These statistical results

[6] BNC data here are modified from Gardner and Davies (2007).

indicate that Korean EFL students exercise fewer PVs than do native speakers and thus support a general claim that the non-native speaker uses fewer PVs than the native speaker does. However, it is not clear yet whether Korean learners' underuse of PVs is due to the strategic behavior of the non-native to avoid using PVs, preferring one-word equivalents (Dagut & Laufer, 1985; Hulstijn & Marchena, 1989; Liao & Fukuya, 2004), or due to the lack of formulaic competence (Alejo, 2010, 2012). Thus, further research is needed with respect to their linguistic preference between one-word verbs and PV equivalents. From a pedagogical perspective, it is worth mentioning that PVs as a grammatical class has a higher frequency than the verb *is* (0.79%), *are* (0.32%), *was* (0.45%), or *were* (0.08%), *have* (0.37%) or *has* (0.13%), or even the total frequency of all above (2.14%). Hence, it can be said that PVs take important part in the composition of Korean EFL students, so English teachers should pay more attention to their students' use of PVs to see in what context the student employs PVs, and whether they use them properly semantically and pragmatically in context since many PVs have more than one meaning (e.g., *make up, turn in*).

With respect to the frequency of LV lemmas, Table 2 demonstrates that Korean EFL students make use of significantly more lemmas in PV constructions (24.9%) than native speakers do (8.0%). To put it in perspective, Korean EFL students use approximately one in every four lemmas in PV constructions (361:90), while native speakers use one in every 12.5 lemmas (19,682:1,572). These raw numbers show that Korean EFL students use on average more lemmas in PV constructions than native speakers do. One way to explain this result might be that the lemmas Korean EFL students used were a relatively small number (361) compared to native speakers (19,682), and they utilized them very frequently. One thing still to be pointed out is that, as some writers point out (e.g., Ansell, 2000), if using PVs rather than their Latinate synonyms (e.g., *use up* vs. *consume*, *gather together* vs. *assemble)* makes the text more colloquial and less formal, then it can be said that Korean EFL students' writing tends to be closer to the spoken language than that of native speakers. Table 3 below presents top 20 of frequently occurring lemmas in PV constructions both in the EEC corpus and in the BNC:

TABLE 3
Frequency of Top 20 LVs in PV Constructions in EEC and BNC

	EEC				BNC		
Lemma	total #	# in PV	% of PV	Lemma	total #	# in PV	% of PV
GO	481	50	10.4	GO	227,103	48,016	21.1
GIVE	228	49	21.5	COME	145,047	36,878	25.4
COME	189	42	22.2	TAKE	173,996	22,970	13.2
GET	516	17	3.3	GET	213,726	20,223	9.5
GROW	33	16	48.5	SET	39,149	18,569	47.4
FIND	154	15	9.7	CARRY	30,572	15,617	51.1
HANG	16	15	93.8	TURN	44,051	13,040	29.6
LOOK	206	13	6.3	BRING	42,567	12,514	29.4
SUM	14	10	71.4	LOOK	109,110	12,226	11.2
LIVE	318	9	2.8	PUT	67,839	11,970	17.6
FALL	24	5	20.8	PICK	14,274	9,997	70
LIE	8	5	62.5	MAKE	210,880	7,368	3.5
PUT	21	5	23.8	POINT	13,767	7,159	52
SET	36	5	13.9	SIT	27,388	7,112	26
SIT	13	5	38.5	FIND	96,010	6,934	7.2
WAKE	5	5	100	GIVE	125,312	6,174	4.9
PICK	8	5	62.5	WORK	63,104	5,985	9.5
MAKE	495	5	1	BREAK	18,642	5,428	29.1
WORK	229	5	2.2	HOLD	46,773	5,403	11.6
BREAK	73	4	5.5	MOVE	37,820	5,197	13.7
total	3067	285	9.3*	total	1,747,130	278,780	24.2*

*: Average of column

As seen in Table 3, the lemma GO is the most frequently used LV in PV constructions in the BNC as well as in the EEC corpus. Besides, GO, COME and GET were in the top four in both corpora. In other words, the most common LVs functioning as the PV in native speakers' corpus are the same as in Korean EFL students' corpus. It seems that Korean EFL students recognize the most frequent English LVs in PV constructions.

There are, however, noticeable differences between the top 20 lemmas in PV constructions in the two corpora. First, among the top 20 lemmas in each corpus, 13 lemmas overlap, but even highly frequent lemmas in PV constructions in the BNC, such as TAKE (22,970 or 13.2%), CARRY (15,617 or 51.1%), TURN (13,040 or 29.6%), BRING (12,514 or 29.4%), etc. do not appear in the top 20 of the EEC corpus. Instead GROW, HANG, SUM, LIVE, FALL, LIE, and WAKE take place in the top 20 list in the EEC corpus, and among these LV lemmas, GROW, HANG, SUM, LIE, and WAKE combine overtly often with AVPs than with any other grammatical class or standing alone. HANG (93.8%) and WAKE (100%) in particular appear with AVPs almost all the times. As such, pedagogically speaking, both the teacher and student should be aware that Korean students are not used to using some of the most common PVs in English, and that some of the lemmas in PV constructions which they exercise lack the combinations with various AVPs.

Second, the overall usage rate of the top 20 lemmas in verb-plus-particle constructions by Korean EFL students (9.3%) is much lower than for native speakers (24.2%). In comparing just the 13 overlapping lemmas, native speakers utilize more of them in PV constructions than Korean EFL students do. To put it another way, Korean EFL students use much less PVs than do native speakers. This result is consistent with the difference between Korean EFL students and native speakers with regard to the total use of PVs seen earlier. Then what specific verb-plus-particle constructions Korean EFL students exercise will be discussed in the following.

Frequency of Phrasal Verbs (PVs) in EEC

As seen above, it seems that Korean EFL students are not hesitant to use PV constructions despite the claim that non-native learners preferably use one-word verbs over PVs when these would be the typical choice by native speakers (Dagut & Laufer, 1985; Hulstijn & Marchena, 1989). The top 60 PVs in the EEC corpus are listed in Appendix A. Table 4 presents the raw count of PV constructions for the top 20 LV lemmas in the EEC corpus:

TABLE 4
AVP Frequency of Top 20 LVs in the PV Construction in EEC

	up	out	back	down	in	over	around	on	by	off	along	through	total
go	4	5	18	1	0	0	3	2	15	0	0	2	50
give	49	0	0	0	0	0	0	0	0	0	0	0	49
come	13	4	22	1	1	0	1	0	0	0	0	0	42
get	5	0	1	1	1	2	0	2	0	2	2	1	17
grow	16	0	0	0	0	0	0	0	0	0	0	0	16
find	0	15	0	0	0	0	0	0	0	0	0	0	15
hang	0	9	0	0	0	2	4	0	0	0	0	0	15
look	2	1	5	0	0	1	4	0	0	0	0	0	13
sum	10	0	0	0	0	0	0	0	0	0	0	0	10
live	0	1	0	0	7	0	1	0	0	0	0	0	9
break	4	2	0	0	0	0	0	0	0	0	0	0	6
fall	0	1	0	2	2	0	0	0	0	0	0	0	5
lie	0	0	0	4	0	0	0	0	0	1	0	0	5
put	1	0	0	1	1	0	0	2	0	0	0	0	5
set	4	0	0	0	0	0	0	0	0	1	0	0	5
sit	0	0	0	4	0	0	1	0	0	0	0	0	5

wake	5	0	0	0	0	0	0	0	0	0	0	0	5
pick	5	0	0	0	0	0	0	0	0	0	0	0	5
make	4	0	0	1	0	0	0	0	0	0	0	0	5
work	0	2	0	0	0	1	0	2	0	0	0	0	5
total	122	40	46	15	12	6	14	8	15	4	2	3	287

A few noticeable phenomena (but intertwined with each other) are observed in the use of verb-plus-particle constructions by Korean EFL students. First, these 20-verb lemmas can be divided into two groups in relation to AVPs; one that a verb combines with various AVPs (e.g., GO + up, out, back, down, around, on, etc.) and another that a verb followed by only one or two particular AVPs, but not others (e.g., GIVE + up, FIND + out). GO, COME, GET, LOOK, and PUT fall in the first group, and the rest of the verbs, GIVE, GROW, HANG, FIND, SUM, LIVE, BREAK, FALL, LIE, SET, SIT, WAKE, PICK, MAKE, and WORK in the second group.[7] As can be seen, despite the high productivity of PVs, Korean EFL students utilize it with only several lemmas, i.e. ones in the first group, constraining most of the others in combining with a certain particle presumably that had been most often heard or seen. For example, GIVE which was one of the most frequent lemmas in PV constructions occurred with *up* alone in the EEC corpus, whereas GIVE combines with *up* 4,186 times, *out* 532, *back* 507, *in* 579 as well as with other AVPs such as *on, off, over* in the BNC although the frequency of *up* is much higher than other particles.

Secondly, Gardner and Davies (2007) found in the BNC that the particle *on* combines often with the verb lemma GO, but never once with the lemma POINT, and the particle *over* combines frequently with the verb TAKE, but never with the verbs SET, POINT, or FIND, and as a result, claim that "the actual verb + particle

[7] The division of the two groups is made by the criteria that 1) when the highest frequency of a lemma combines with any particular particle alone exceeded the total frequency of the lemma combine with the rest of particles, and 2) when a lemma combined with fewer than 3 particles, the lemma falls in the second group. Otherwise, it is in the first group.

combinations are highly idiosyncratic" (p. 349). When compared to the BNC, however, the verb + particle combinations that such verbs like GIVE, GROW, FIND, etc. in the second group in the EEC corpus demonstrate do not correspond to the idiosyncratic nature of particles which Gardner and Davies assert. For example, *down* is the most frequent particle that combines with the verb BREAK in the BNC, yet, it never combines with BREAK in the EEC corpus. Moreover, as seen in Table 4, there is a lot of zero that indicates no occurrence at all. For example, the particle *over* never combines with COME in the EEC corpus, while combining with it a total of 1,004 times in the BNC, and the particle *out* never combines with GET in the EEC corpus, while combining with it 3,545 times in the BNC. Hence, it can be said that if PV constructions are idiosyncratic in the actual use of English, as Gardner and Davies (2007) claim, then that should be explicitly taught to non-native speakers.

Finally, a closer look at the differential attributions of AVPs in the two groups of verb lemmas and the comparison between the EEC and BNC indentify a certain influence of English in use on Korean EFL students' writing. The verbs in the first group, which combine with various AVPs in the EEC corpus, happen to be those whose frequency attributions of different AVPs are not considerably dissimilar in the BNC. In other words, English lexical verbs that combine with various AVPs at a fairly similar frequency also combine with various particles in Korean EFL students' writing. For example, in the case of GET in the BNC, it combines with *out* 3,545 times, *up* 3,936, *on* 2,696, *back* 4,552, *down* 1,538, etc., and Korean students employ it combining with *up, back, down, in, over, on*, etc. The first group of lemmas in the EEC corpus, i.e., GO, COME, LOOK, and PUT, all belongs to this category. It appears that Korean EFL students perceive frequent verb + particle combinations used by native speakers.

Furthermore, where a particular verb + particle combination accounts for the vast majority of cases in the BNC, it is found that Korean learners have adopted that combination as the only PV to involve that verb. For example, FIND + out and PICK + up are the dominant verb + particle combinations for these lemmas in the BNC, as other combinations (e.g. FIND + up, on, etc. PICK + out, on, etc.) rarely occur. That is to say, FIND combines with *out* 6,619 times, *up* 33, *on* 9, *back* 128,

etc; PICK combines with *up* 9,037 times, *out* 856, *on* 35, *back* 3, etc. Korean EFL students tend to adopt just these two combinations and completely ignore the less frequent ones. The pattern is similar with most of the second group lemmas, such as GIVE, SET, SIT, and MAKE.

This finding seems to require further scrutiny as to over- and under-use of PVs by non-native learners; is it a product of input frequency, or caused by L1 influence or strategic avoidance of learners, as some studies argue (Dagut & Laufer, 1985; Hulstijn & Marchena, 1989), or something else?

BREAK and WORK are exceptions, however. They do not fit in either of the categories. According to the criteria of the group division, the two verbs belong to the second group, combining with fewer than three different particles. The BNC analysis also shows that WORK behaves like the verbs in the second group, manifesting the highest frequency of the WORK + out combination exceeding the total frequency of the rest of combinations, i.e., *out* 4,703 times, *up* 334, *on* 411, *back* 36, *down* 98, etc. WORK in the EEC corpus, however, does not combine noticeably more often with *out* (2) than *on* (2) or *over* (1). With respect to BREAK, it combines with various particles at arguably similar frequencies, i.e., *out* 996 times, *up* 1,286, *down* 2,199, etc. in the BNC, showing the BREAK + down combination the highest frequency. This feature puts BREAK in the first group. It, however, combines exclusively with *up* and *out* in the EEC corpus, not revealing even once the most frequent combination (BREAK + down) in the BNC, thus falling itself in the second group.

Overall, Korean EFL students seem to acknowledge frequently used verb-plus-particle combinations by native speakers. Yet, of pedagogical concern is the fact that there are some combinations that reveal the differentiation between native speakers and Korean EFL students in terms of their frequency, such as WORK, BREAK, and other lemmas which do not appear in Korean learners' writing, but are very frequent in the BNC. This certainly needs to be recognized by both the English teacher and student.

CONCLUSION

Taking the perspective provided by a usage-based approach to language, this study attempted to identify the most frequent verbs and adverbial particles in verb-plus-particle constructions and PVs in Korean EFL students' writing and then compare the results to the actual use of native speakers in the BNC, aiming to provide a better understanding of Korean learners of English. The findings include that *up, out, back*, and *down* were the most frequent adverbial particles in the EEC corpus, and it was the same for the BNC. However, native speakers used more various AVPs (16) than the Korean learners did (12), and the particles *on* and *about,* which were used frequently as AVPs in the BNC, were seldom utilized in the EEC corpus.

Korean EFL students made use of fewer PVs (2.9%) than did native speakers (5.0%); yet, with respect to the average usage rate of lemmas in PV constructions, they used significantly more lemmas in PV constructions (24.9%) than native speakers did (8.0%). This seems to have resulted from a relatively small number (361) of lemmas Korean students employed, compared to native speakers (19,682). From the limited number, they frequently utilized the verbs in PV constructions.

Among the top 20 lemmas in PV construction in the BNC and EEC corpus, 13 lemmas overlapped, but even highly frequent lemmas in PV constructions in the BNC, such as TAKE (22,970), CARRY (15,617), TURN (13,040), BRING (12,514), etc. did not appear in the top 20 in the EEC corpus. The most common LVs, e.g. GO, COME and GET, functioning as the PV in the native speakers' corpus, however, were the same as in the Korean EFL students' corpus.

The top 20-verb lemmas in the EEC corpus were divided into two groups in relation to the combinations with AVPs; a group of verbs that combined with various AVPs and a group of verbs that were followed by only one or two particular AVPs, but not others. The verbs in the first group in the EEC corpus revealed a fairly similar frequency throughout all disparate AVPs in the BNC. Thus, it can be argued that frequent combinations used by the native speaker were recognized and utilized by Korean EFL students. As for the verbs in the second group in the EEC corpus, Korean students used only certain verb + particle combinations which

accounted for the vast majority of cases in comparison with other combinations in the BNC. This statistical comparison between the BNC and EEC corpus seemed to show some relationship between the frequency of PVs in native speakers' use and non-native learners' use of PVs.

Although overall results revealed that highly frequent PVs in native speakers' use appeared to influence Korean students' use of them, there were some exceptional cases that both the English teacher and student alike should take account. For example, some frequent verbs (e.g., TAKE, CARRY, TURN) in the BNC did not appear in the top 20 list of the EEC corpus, and some very frequent combinations like 'BREAK + down' in the BNC were not used by Korean learners.

Further research is needed to investigate, in a more theoretical sense, which types of PVs (literal or figurative) Korean learners of English more often use and in what context they prefer using PVs or one-word verbs, and in a more practical sense, whether they use PVs properly semantically and pragmatically in the context and what we as English teachers can do to help the student become more native-like in terms of the use of verb-plus-particle constructions.

THE AUTHOR

Mi-Lim Ryoo is an associate professor of English Language and Literature at Korea Maritime University in South Korea. Her research interests include second language writing, L2 learner corpus studies, and ESP.
Email: ryoo@hhu.ac.kr

REFERENCES

Alejo, R. (2010). Making sense of phrasal verbs: A cognitive linguistic account of L2 learning. *AILA Review 23*, 50-71.

Alejo, R. (2012). Variability in L2 acquisition of formulaic sequences: A study of linguistic factors in the use and learning of phrasal verbs by non-native speakers. *Language Value*, *4*(1), 33-62.

Altenberg, B. (1998). On the phraseology of spoken English: the evidence of recurrent word-combinations. In A. P. Cowie (Ed.), *Phraseology: Theory, analysis and applications* (pp. 101-122). Oxford: Clarendon Press.

An, K- S. (2012). Acquisition of semantic categories of English phrasal verbs by Korean speakers. *English Language and Literature, 25*(3), 343-369.

Ansell, M. (2000). English grammar: Explanations and exercises. Retrieved March 28, 2011, from http://www.ansell-uebersetzungen.com/gramdex.html.

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

Bolinger, D. (1971). *The phrasal verb in English.* Cambridge: Harvard University Press.

Celce-Murcia, M., & Larsen-Freeman, D. (1999). *The grammar book. An ESL/EFL teacher's course*. Boston: Heinle & Heinle.

Cobb, T. 2003. Analyzing late interlanguage with learner corpora: Quebec replication of three European studies. *Canadian Modern Language Review, 59*, 393-423.

Cowie, A. P. (1994). Phraseology. In R. E. Asher (Ed.), *The encyclopedia of language and linguistics* (pp. 3168-3171). Oxford: Pergamon Press.

Cowie, A. P. (1998). Phraseological dictionaries: some east-west comparisons. In A. P. Cowi (Ed.), *Phraseology: Theory, analysis, and application 3* (pp. 209-228). Oxford: Clarendon Press.

Dagut, M., & Laufer, B. (1985). Avoidance of phrasal verbs: A case for contrastive analysis. *Studies in Second Language Acquisition, 7,* 73-79.

Darwin, C. M., & Gray, L. S. (1999). Going after the phrasal verb: An alternative approach to classification. *TESOL Quarterly*, 33, 65-83.

Dehé, N. (2002). *Particle verbs in English: Syntax, information structure and intonation.* Amsterdam: John Benjamins.

Durrant, P., & Schmitt, N. (2009). To what extent do native and non-native writers make use of collocations. *IRAL, 47*, 157-177.

Ellis, N. C. (2008). Phraseology: The periphery and the heart of language. In F. Meunier & S. Granger (Eds.), *Phraseology in foreign language learning and teaching* (pp. 1–14). Amsterdam: John Benjamins.

Fletcher, B. (2005). Register and phrasal verbs. *MED Magazine. The Monthly Webzine of McMillan English Dictionaries, 33*, 45-54.

Francis, A. K. (1958). *The structure of American English*. New York: Ronald Press.

Frank, T. (1989). On lexical innovation. *Textus, 2*, 113-140.

Fraser, B. (1976). *The verb-particle combination in English*. New York: Academic Press.

Gardner, D., & Davies, M. (2007). Pointing out frequent phrasal verbs: A corpus-based analysis. *TESOL Quarterly, 41*, 339-359.

Gass, S. M., & Selinker L. (1992). *Language transfer in language learning*. Philadelphia: John Benjamins.

Goldberg, A. (2009). The nature of generalization in language. *Cognitive Linguistics, 20*, 93–127.

Granger, S., & Meunier, F. (2008). *Phraseology: An interdisciplinary perspective*. Amsterdam: John Benjamins.

Granger, S., & Paquot, M. (2008). Disentangling the phraseological web. In S. Granger & F. Meunier (Eds.), *Phraseology: An interdisciplinary perspective* (pp. 27–50). Amsterdam: John Benjamins.

Gries, S. (2008). Phraseology and linguistic theory: A brief survey. In S. Granger & F. Meunier (Eds.), *Phraseology: An interdisciplinary perspective* (pp. 3-25). Amsterdam: John Benjamins.

Howarth, P. (1998). Phraseology and second language proficiency. *Applied Linguistics, 19*, 24-44.

Hulstijn, J. H., & Marchena, E. (1989). Avoidance: Grammatical or semantic causes? *Studies in Second Language Acquisition, 11,* 241-255.

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

Kennedy, G. (2008). Phraseology and language pedagogy. In F. Meunier & S. Granger (Eds.), *Phraseology in foreign language learning and teaching* (pp. 21–41). Amsterdam: John Benjamins.

Laufer, B., & Eliasson, S. (1993). What causes avoidance in L2 learning: L1-L2 difference, L1-L2 similarity, or L2 complexity? *Studies in Second Language Acquisition, 15*, 35-48.

Liao, Y., & Fukuya, Y. J. (2004). Avoidance of phrasal verbs: The case of Chinese learners of English. *Language Learning, 54*, 193-226.

McArthur, T. (1989). The long-neglected phrasal verbs. *English Today, 5*(2), 38-44.

McCarthy, M., & O'Dell, F. (2004). *English phrasal verbs in use.* Cambridge: Cambridge University Press.

Meyer, G. A. (1975). *The two-word verb: A dictionary of the verb-preposition phrases in American English*. Hague: Mouton.

Moon, R. (1998). *Fixed expressions and idioms in English*. Oxford: Clarendon Press.

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

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

O'Dowd, E. M. (1998). *Preposition and particles in English: A discourse-functional account*. Oxford: Oxford University Press.

O'Keeffee, A., McCarthy, M., & Carter, R. (2007). *From corpus to classroom language use and language teaching*. Cambridge: Cambridge University Press.

Perdek, M. (2010). *Getting through to phrasal verbs: A cognitive organization of phrasal verb entries in monolingual pedagogical dictionaries of English*. Retrieved December 15, 2011 from http://www.euralex.org/elx_proceedings/Euralex2010/

Quirk, R., Greenbaum, S., Leech G., & Svartvik, J. (1985). *A comprehensive grammar of the English language*. London: Longman.

Römer, U. (2009). The inseparability of lexis and grammar: Corpus linguistic perspectives, *Annual Review of Cognitive Linguistics, 7,* 141–163.

Scott, M. (2008). *WordSmith Tools version 5*. Liverpool: Lexical Analysis Software.

Sinclair, J. (1991). *Corpus, concordance and collocation*. Oxford: Oxford University Press.

Sinclair, J. (1996). The search for units of meaning. *Textus, 9,* 75–106.

Sinclair, J. (2004). *Trust the text*. London: Routledge.

Sinclair, J. (2008). The phrase, whole phrase, nothing but phrase. In S. Granger & F. Meunier (Eds.), *Phraseology: An interdisciplinary perspective* (pp. 407-410). Amsterdam: John Benjamins.

Sroka, K. (1972). *The syntax of English phrasal verbs.* Hague: Mouton.

Stubbs, M. (2001). *Words and phrases: Corpus studies of lexical semantics*. Oxford: Blackwell.

Stubbs, M. (2009). Technology and phraseology: With notes on the history of corpus linguistics. In U. Römer & R. Schulze (Eds.), *Exploring the lexis-grammar interface* (pp. 15-31). Amsterdam: Benjamins.

Sung, I- H. (2012). On the productivity of word-formation of English phrasal verbs through corpus analysis. *Studies in English Language & Literature, 38*(1), 201-221.

Tyler, A. (2010). Usage-based approaches to language and their applications to second language learning, *Annual Review of Applied Linguistics, 30,* 270–291.

Waibel, B. (2007). *Phrasal verbs in learner English: A corpus-based study of German and Italian students.* Unpublished Ph.D. Dissertation, Albert-Ludwigs-Universität, Freiburg, Germany.

Wray, A. (2002). *Formulaic language and the lexicon*. Cambridge: Cambridge University Press.

Zyzik, E. (2009). The role of input revisited: Nativist versus usage-based models. *L2 Journal*, *1*(1), 42-61.

APPENDIX A

Frequency of Top 60 PV lemmas in EEC

Rank	PV	frequency	Rank	PV	frequency
1	give up	49	26	speak up	3
2	come back	22	26	write down	3
3	go back	18	33	break out	2
4	grow up	16	33	clean up	2
5	go on	15	33	cry out	2
5	go by	15	33	cut down	2
5	find out	15	33	end up	2
8	come up	13	33	fall down	2
9	sum up	10	33	fall in	2
10	hang out	9	33	follow up	2
11	live in	7	33	get along	2
12	go out	5	33	get off	2
12	get up	5	33	get on	2
12	look back	5	33	get over	2
12	wake up	5	33	get through	2
12	pick up	5	33	go through	2
17	go up	4	33	keep on	2
17	come out	4	33	look up	3
17	hang around	4	33	pop up	2
17	look around	4	33	pour out	2
17	lie down	4	33	raise up	2
17	set up	4	33	shine up	2

17	sit down	4	33	show off	2
17	make up	4	33	stress out	2
17	break up	4	33	take off	2
26	carry out	3	33	talk over	2
26	catch up	3	33	turn off	2
26	cheer up	3	33	work on	2
26	figure out	3	33	work out	2
26	shout out	3	33	worry about	2