English Language & Literature Teaching, Vol. 23, No. 1 Spring 2017

Korean High School Learners' Knowledge of Adjective-Noun and Verb-Noun Collocations by L1-L2 Congruency and L2 Proficiency

Eun-Sun Kim[*]
(Hanyang University)
Yuah V. Chon[**]
(Hanyang University)

Kim, Eun-Sun & Chon, Yuah V. (2017). Korean high school learners' knowledge of adjective-noun and verb-noun collocations by L1-L2 congruency and L2 proficiency. *English Language & Literature Teaching*, *23*(1), 77-99.

While single word items have more often been the unit of interest for second language (L2) vocabulary learning, there has been less attention paid to how L2 learners may demonstrate knowledge of collocations. L2 high school learners were asked to retrieve collocations according to different degrees of congruency (i.e., incongruent vs. congruent) that are shared between L1 and L2 collocations. For this purpose, a receptive collocation test (RCT), counterbalanced for incongruent and congruent collocations, and adjective-noun (AN) and verb-noun (VN) collocations, was developed by matching normalized frequencies of the collocation items from the Corpus of Contemporary American English (COCA). The 133 high school sophomores were found to have more problems with the incongruent collocations than with the congruent items. However, correct responses were more frequent for the incongruent VN collocations than those of the congruent VN collocations. The results demonstrate that synonymous collocates can also be the source of problems in learning collocations. The study also empirically demonstrates that L2 proficiency is the foundation for development of collocational knowledge. Implications on teaching collocations are presented in light of the obtained results.

[collocations/congruency/proficiency/collocates/COCA]

[*] First author
[**] Corresponding author

I. INTRODUCTION

The L2 learners' knowledge of vocabulary is a fundamental skill regardless of the language skill that is involved. However, more recent research on learning vocabulary has shifted its focus from single word items to noticing the importance of multiword items, such as collocations. A number of studies have mentioned that formulaic language plays an important role in language processing, production, and acquisition (Chon & Lee, 2015; Erman & Warren, 2000; Conklin & Schmitt, 2008; Kim, 2010; Wray, 2008). Nation (2001) has mentioned that knowing a word includes not only knowing its form and meaning, but also being able to retrieve them in sequence or chunks. Conzett (2000) has suggested that knowing collocations will enable learners to make accurate use of the words they know, and collocation proficiency is eventually expected to contribute to native-like writing, fluent reading and faster comprehension (Conklin & Schmitt, 2008; Martinez & Murphy, 2011). However, the teaching of vocabulary, to a large extent in the Korean school secondary context, is still limited to teaching and learning single word items (Kim, 2005). As such, when learners are asked to combine words in discourse, this may become the source of a problem in communication (Hill, 2000). For instance, Korean learners may know *last, final, decision, take, exam, have* and *party,* but they would have problems retrieving *final decision* and instead produce **last decision* (Note: * = unacceptable). For *take an exam*, learners may retrieve **see an exam* instead; **open a party* may be retrieved rather than *have a party*. The result is that learners will combine words in a haphazard manner, often producing unacceptable collocations. This is expected to occur when there are mismatches between the constituent words of L1 and L2 collocations. These will be referred to as incongruent collocations. On the other hand, the knowledge of L1 collocations may help learners retrieve L2 collocations when the constituent words share semantically similar words. These would be categorized as congruent collocations. However, the cross-linguistic explanation may not be so simple when word confusion problems occur due to learners retrieving synonymous collocates (e.g., *final/last*) in the process of trying to find the best matching collocate for a node word (e.g., *decision*) (Jung, Sur, & Kim, 2007; Jung, Park, & Kim, 2010; Kim, 2012).

By recognizing the importance of collocations and realizing that they can cause problems for L2 Korean learners of English, the purpose of the present study was to examine how L2 learners performed when asked to retrieve congruent (i.e., having an equivalent first language construction) collocations and incongruent (i.e., not having an equivalent L1 construction) collocations. Another was to examine how their collocation knowledge had a relationship to the learners' L2 language proficiency.

II. BACKGROUND

1. Learning of L2 Collocations

While multiword items may include 'collocations' (Boers, Eyckmans, Kappel, Stengers, & Demecheleer, 2006; Nation, 2001), others are idioms (e.g., *kick the bucket*), fixed expressions (e.g., *to and fro*), fillers (e.g., *sort of*), functional expressions (e.g., *excuse me*), and standardized phrases (e.g., *as a matter of fact*). According to Liu (2010), collocation means "primarily habitual combinations of words, such as *do* (not *make*) *laundry* and *make* (not *do*) *a case"* (p. 5). Benson (1989) contends that "collocations should be defined not just as 'recurrent word combinations,' but as 'arbitrary recurrent word combinations'" (p. 3). Similarly, Lewis (2000) also sees collocations as being arbitrary: *high/tall building*, *tall boy*, but not **high boy*. *Prices rise and fall*; you can *rise to the occasion* but not **fall to the occasion*" (p. 26). As such, one of the problems in learning collocations is that simple rules cannot be applied in acquiring them (Yamashita & Jiang, 2010). Nor will the knowledge of single lexical items in a collocation result in the automatic acquisition of lexical combinations (Wolter, 2006).

Boers, Lindstromberg, and Eyckmans (2014) draw our attention to some of the factors that may cause difficulty for learning L2 collocations. They reason that it is the adult readers' greater attention to single words rather than recurring word combinations that will slow down their acquisition of collocations. Another possible factor is explained by salience. Learners are unlikely to be struck by novelty when the content words that make up a given collocation are likely to be quite common (e.g., *make* + *mistake* and *tall* + *tree*). A third explanation is that it is common for one of the constituents of a collocation to make less of a contribution to the collocation's meaning. This is most notable for verb-noun collocations where the "verb is a high-frequency, multi-purpose item (e.g., *have, make,* or *do*), and it is the noun rather than the verb that carries most of the meaning (e.g., *have a nightmare, make a mistake,* or *do business*)" (p. 46).

As another aspect of learning collocations, the cross-linguistic nature thereof will make learning subject to first language (L1) influence. It is this specific area of learning collocations, that is, on how congruency may have an effect on learning collocations that the current research intends to explore.

2. Congruency and L2 Collocational Knowledge

Transfer from L1 is a well-attested phenomenon in L2 collocation use (Jiang, 2004; Kroll & Stewart, 1994; Laufer & Girsai, 2008; Nesselhauf, 2005). In trying to understand how L2 collocations are processed, studies in the area of L2/bilingual mental

lexicon have demonstrated that L2 lexical items are processed via mediation of the L1 lexicon, even at advanced levels of L2 proficiency (Wolter, 2006; Wolter & Gyllstad, 2011; Yamashita & Jiang, 2010).

Regarding the influence of L1 lexical/conceptual knowledge on L2 lexical knowledge, Wolter (2006) noted that the SLA research literature was not adequately addressing this. He presented a theoretical account of how learners might draw upon L1 lexical and conceptual knowledge to make assumptions about connections between words in the L2 lexicon. He explains that the conceptual connection between *office* and *room* is basically the same for any language user, so the conceptual connection will basically be the same for anyone. On the other hand, when there is a significant difference between L1 and L2 lexical items in the representation of a specific concept, he claimed that the lexical network is likely to undergo fundamental restructuring in order to lexically or conceptually modify its existing mental lexicon.

In a similar vein, Yamashita and Jiang (2010) investigated first language (LI) influence on the acquisition of second language (L2) collocations by comparing the performance on a phrase-acceptability judgment task among native speakers of English, Japanese English as a second language (ESL) users, and Japanese English as a foreign language (EFL) learners. The researchers were able to ascertain that (a) both LI congruency and L2 exposure affect the acquisition of L2 collocations; (b) it is difficult to acquire incongruent collocations even with a considerable amount of exposure to L2; and (c) once stored in memory, L2 collocations are processed independently of LI.

Wolter and Gyllstad (2011) also investigated the effects of congruency and incongruency on collocational processing. They used a primed lexical decision task to investigate whether or not the presentation of a verb prime was more likely to facilitate faster recognition of a (object) noun target if the resulting verb + (object) noun collocation was congruent versus incongruent. Participants included NSs of English as well as advanced NNSs of English (i.e., L1 Swedish). The results showed for the NNSs that they responded significantly faster to the congruent items over the incongruent items. With Korean EFL learners, Kim (2012) investigated the influence of L1 on 111 Korean high school learners' knowledge of 16 adjective-noun and 18 verb-noun English collocations. While the collocation tests involved both congruent and incongruent collocations, she found synonymy affecting the acquisition of congruent collocations when the options provided caused word confusion problems. Nevertheless, the incongruent collocations were found to be more difficult to acquire than the congruent collocations.

3. Collocations and L2 Proficiency

Since the ultimate goal of learning L2 collocations is to help learners become more

fluent and accurate in their knowledge of a second language, it was felt important to seek the relationship between the knowledge of L2 collocations and L2 proficiency. Howarth (1998) found from the comparison of NS performance and NNS errors that advanced level learners are lexically competent with having successfully internalized the more restricted collocations. The knowledge of L2 collocations has also been studied in relation to learners' speaking proficiency. Boers et al., (2006) report on how the use of formulaic sequences can help language learners come across as proficient speakers through an L2 interview. In the area of L2 writing, Alternberg and Granger (2001) examine EFL learners' use of high-frequency verbs, in particular with the use of the verb 'make.' Results showed that EFL learners, even at high proficiency levels, were having great difficulty with phrasing expressions that included 'make.' Nizonkiza (2011) assessed the relationship between EFL proficiency, lexical competence, and collocational competence. Kim (2003) also found with ninety second year high school learners in a multiple regression analysis that knowledge of collocations was a significant predictor of English proficiency. The measures of L2 speaking and writing in the studies have their advantages for being domain-specific. However, the studies have not been able to pay systematic attention to categorizing the learners by L2 proficiency groups.

While the previous studies shed light on how learners may find a processing advantage for congruent collocations over incongruent collocations, the studies may fall short by taking a simplistic view of how L2 learners respond to the collocations when tested for effects of congruency and incongruency. That is, previous studies have not been able to control the frequency rate of collocations, which may have an effect on the learners' knowledge of collocations, but we tried to control this factor in the study. Also, research on L2 collocational competence also calls for more examination on the relationship between collocational proficiency and L2 general proficiency of learners with the same L1 background by valid measures of proficiency. For instance, Yamashita and Jiang (2010) grouped their participants into English native speaker, ESL, and EFL learners. However, these categories do not necessarily reflect the proficiency levels of the ESL and EFL learners since there may be range of proficiency profiles within a group. While the primary goal of learning English is to develop L2 general proficiency, the researchers empirically sought to investigate if developing L2 general proficiency is related to improving L2 collocational knowledge. This is because L2 collocations do not receive the same level of attention as single word items in L2 learning. Knowing that EFL Korean high school learners have a low sense of awareness on how vocabulary learning can occur with collocations, the following research questions guided us through the study:

1. How does the Korean EFL high school learners' knowledge of collocations differ

according to congruency (i.e., Incongruent Collocations vs. Congruent Collocations) and parts of speech (i.e., AN and VN)?

2. What is the Korean EFL learners' knowledge of collocations according to L2 proficiency?

III. METHOD

1. Participants

For the purpose of the present study, 160 EFL learners were recruited from a private boy's high school in the Northern Jeolla Province in South Korea. The learners in the present study, who were in their second year, had studied English since the 3rd year of elementary school. At the time of the study, the learners were receiving three to four hours of English instruction per week. None of the learners reported on having lived in an English-speaking country for study purposes. Of the learners, 94% of the learners replied that they did not know about collocations or either reported that they had not associated vocabulary learning with learning collocations. Due to the aim of the study which was to accurately assess the learners' knowledge of collocations, learners who had reported 'Don't know' for more than half of the items in the collocation test were excluded from the actual data analysis, and this left the researchers with 133 learners to analyze. When the 133 learners' scores of the College Scholastic Ability Test (CSAT hereafter) from a previous mock test was calculated, their scores ranged from 15 (min) to 98 (max) totaling a mean of 51.12 points (*SD* = 19.73). When the learners' stanine levels were examined to determine their general English language proficiency, the levels could be categorized as 'Stanine levels 1-3,' 'Stanine levels 4-6,' and 'Stanine levels 7-9' (See Table 1).

TABLE 1
Participants of the Study and Stanine Levels

Stanine Levels	*N*	Mean	*SD*
Levels 1-3	17	85.82	10.41
Levels 4-6	86	52.71	11.04
Levels 7-9	30	26.90	5.79

2. Instruments and Materials

1) Receptive Collocation Test

As the purpose of the current study was to assess the collocation proficiency of the learners, a receptive collocation test (RCT hereafter) was designed. Although there may be several types of collocations (Benson, 1997), those that may be considered most problematic to L2 learners were selected for the current study, that is, 'adjective + noun' and 'verb + noun' collocations (Al-Zahrani, 1998; Chen, 2002; Liu, 1999; Webb, Newton, & Chang, 2013). Also, *English Collocations in Use* (McCarthy & O'Dell, 2005) and previous studies on the testing of collocations with Korean learners of English were referenced for selecting collocation items (Jung et al., 2007; Jung et al., 2010; Kim, 2012; Ko, 2010; Lee & Kim, 2005). Most of all, collocations were checked manually in the *Oxford Collocations Dictionary* (OCD hereafter) to see if they could be treated as a unit thereof.

Since one of the interests of the study was to inquire how congruency would have an effect on the learners' knowledge of collocations, it was necessary to identify the collocations as being either congruent or incongruent. For this purpose, the *Oxford Advanced Learner's English-Korean Dictionary* (2008) and the *Doosan Dong-A Prime English-Korean Dictionary* (2011) were used to judge if the collocation items should be categorized as either congruent or incongruent according to the transparency of L1 translations in relation to L2. As a result, 86 collocation items could be identified for testing, and these items were categorized into four types: congruent adjective-noun (CAN) (23 items), incongruent adjective-noun (IAN) (23 items), congruent verb-noun (CVN) (20 items), and incongruent verb-noun (IVN) (20 items) collocations.

In the process of developing the RCT, information on frequency per million (FPM) was also referenced in order to control the potential exposure level between the congruent and incongruent collocations. For this, FPM was utilized from the Corpus of Contemporary American English (Davies, 2008-2012, hereafter COCA) to counterbalance the selection of collocations according to congruency. That is, FPM for CAN and IAN were respectively 0.67 and 0.34, and FPM for CVN and IVN were respectively 0.54 and 0.80 so that the differences in frequency were only a decimal difference. Differences were also not significant ($p > .05$) between the collocations within the same part of speech. To search for the FPM of the adjective-noun collocations, the 'noun' was set as the node while the 'adjective' was treated as the collocate with the horizon set at L1 and R0. For the verb-noun collocations, the 'noun' was set as the node while the 'verb' was treated as the collocate with the horizon set at L2 and R0. An illustration of the search for the adjective-noun collocations is presented

in Figure 1.

Another procedure of the test involved including exemplary sentences that would prompt the use of collocations, and these were selected from the OCD and COCA. The sentences were kept syntactically simple so that learners could concentrate on solving the collocation items. The items were developed for multiple-choice format, and thus the options were developed accordingly for the learners to select; these consisted of a target collocate, a distracter, and 'Don't know,' which was to reduce the number of random guesses that learners would make. When any of the words in a given sentence was deemed difficult for the Korean EFL learners at the high school level, their definitions were provided in L1. A sample collocation testing item is presented as follows (See Figure 2).

FIGURE 1
Analysis of Collocation Frequency (per Million) at COCA

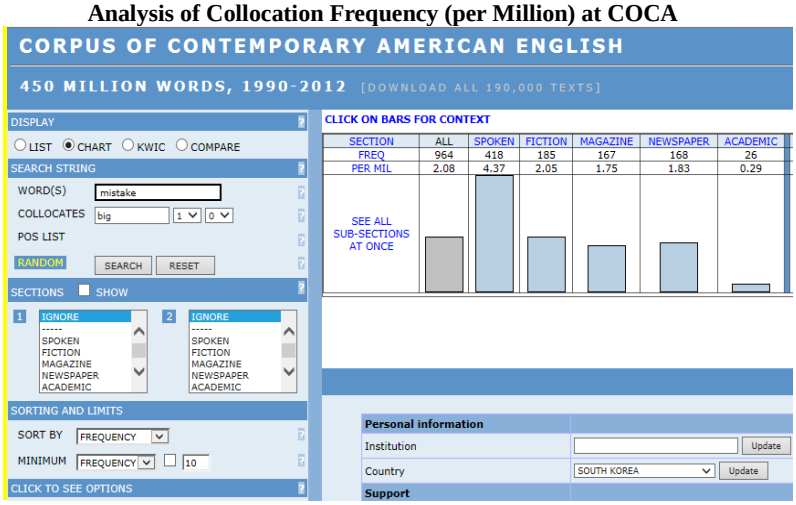

FIGURE 2
A Sample Collocation Item

32. A [ ] **sea** after the storm glistened in the sunlight. *glisten 반짝이다
① quiet ② calm ③ 잘 모르겠다 (Don't know)

2) Mock College Scholastic Ability Test (CSAT)

The learners' proficiency levels were assessed with a mock test of the Korean CSAT (i.e., *jeong-guk-yunhap-pyung-ga*), which had been conducted in March of 2016 throughout the country by the Seoul Metropolitan Office of Education. There are a total

of 45 items on the test and the items add up to a score of 100. There are sections for listening (26 points), speaking (11 points), reading (41 points), and writing (22 points). As such, the instrument was considered valid for assessing the learners' L2 general language proficiency.

3. Data Analysis

All quantitative results were analyzed with SPSS 21.0. To calculate the results of the RCT, each collocation was entered in SPSS as either '1' when the item was correct and '0' when the response was incorrect; any responses with 'Don't know' were also treated as '0.' For research question (RQ) 1, the responses of the RCT were calculated for descriptive statistics (e.g., mean and SD), and inferential statistics with the paired t-test were conducted to check for any differences in the learners' responses for congruent and incongruent collocations (i.e., AN and VN collocations, CAN and IAN collocations; and CVN and IVN collocations). For RQ 2, one-way ANOVA was conducted to ascertain the differences according to the learners' proficiency groups with the learners' responses on the RCT set as the dependent variable.

4. Procedure

The RCT was administered during the learners' regular English sessions under the supervision of the classroom teacher. For the learners to solve the 86 RCT items, 40 minutes was provided. However, more time was provided upon request. Test takers were given both oral and written instructions. They were also led through a sample item to ensure familiarity with the test. The learners were reminded to check 'Don't know' when not sure about the most appropriate collocate rather than to make random guesses.

IV. RESULTS AND DISCUSSION

1. Learners' Knowledge of Collocations by Congruency

This section reports on the results of RQ 1. As a whole, while the learners' demonstrated greater knowledge of congruent collocations (M = 0.44, SD = 0.18) than incongruent collocations (M = 0.41, SD = 0.18), the difference was significant (t = 3.33, df = 132, p = .001). This is likely due to the fact that constituent words of congruent collocations share common meaning associations between L1 and L2. In contrast, the difference in the constituent words of the L1 and L2 collocations would have caused

confusion for the learners when tasked with finding the most appropriate collocate for the collocations (Jiang, 2004; Kim, 2012; Wolter & Gyllstad, 2011; Yamashita & Jiang, 2010).

Regarding different parts of speech, the learners' demonstrated better knowledge of VN collocations (M = 0.43, SD = 0.21) than AN collocations (M = 0.41, SD = 0.16), and this difference was significant (t = -2.064, df = 132, p = .041). However, as seen in Table 2, the learners demonstrated different levels of collocation knowledge according to parts of speech. Learners demonstrated greater knowledge of CAN collocations than of IAN collocations (p = .000), whereas the learners' knowledge of IVN collocations was greater than that of CVN collocations (p = .010). The different patterns that emerged for the AN and VN collocations warranted an explanation through an examination of the mean values of the collocation items from the RCT.

TABLE 2

'Adj + Noun' and 'Verb + Noun' Collocations by Congruency

Part of Speech	Congruency	Mean	*SD*	*t*	*df*	*Sig.*
Adj+Noun	Congruent	0.46	0.17	7.33	132	.000**
	Incongruent	0.36				
Verb+Noun	Congruent	0.42	0.22	-2.63	132	.010*
	Incongruent	0.45				

Note: ** p < .01, * p < .05

For AN, as seen in Table 3, those items that recorded a mean below 0.3 for CAN were *major storm*, *severe pain*, *true colors*, *clear skin*, and *clean election* (5 items), whereas *major surgery*, *thick accent*, *rich taste*, *fair wind*, *slight smile*, *heavy scent*, *heavy seas*, *slim chance*, and *ugly truth* (9 items) did not reach a mean of 0.3 for IAN. This may indicate that incongruent collocations are more challenging for the learners due to mismatches of the constituent words between the L1 and L2 collocations. In the process of solving the items, the learners' overreliance on L1 translations may have contributed to the lower mean responses of IAN. For instance, when learners in the study were asked to retrieve the IAN *slim chance*, the learners seem to have retrieved an L1 collocation *heebakhan ganeungsung* (i.e., 희박한 가능성, **rare chance*) when promoted by the node 'chance' to retrieve 'slim' in the RCT. As a result, there were more learners who chose **rare chance* (83.5%) over *slim chance* (16.5%).

TABLE 3
Mean Responses for 'Adj + Noun' Collocations

Congruent Collocations			Incongruent Collocations		
Adj+Noun	*M*	*SD*	Adj+Noun	*M*	*SD*
high/tall* position	0.88	0.33	much*/heavy rain	0.67	0.47
large*/big mistake	0.66	0.47	heavy*/deep fear	0.57	0.50
light*/bright smile	0.65	0.48	quiet*/calm sea	0.51	0.50
nice*/good relations	0.62	0.49	thin*/faint smell	0.50	0.50
close*/near future	0.57	0.50	big*/deep sorrow	0.50	0.50
false*/fake smile	0.56	0.50	thick/strong* fog	0.46	0.50
born*/natural abilities	0.54	0.50	rare*/thin air	0.44	0.50
strong*/tough meat	0.53	0.50	easy*/light exercise	0.44	0.50
broad*/wide network	0.52	0.50	strong*/heavy drinker	0.44	0.50
fast*/quick shower	0.50	0.50	near*/short distance	0.41	0.49
strong*/powerful tool	0.50	0.50	light*/small accident	0.40	0.49
nice*/good guess	0.47	0.50	empty*/bare branches	0.35	0.48
hot*/heated argument	0.46	0.50	light*/minor injury	0.33	0.47
naked*/bare feet	0.45	0.50	crowded*/heavy traffic	0.30	0.46
last*/final decision	0.39	0.49	big*/major surgery	0.29	0.46
tough*/rough surface	0.31	0.46	tough*/thick accent	0.29	0.45
comfortable*/relaxed mood	0.30	0.46	deep*/rich taste	0.26	0.44
tough*/rough voice	0.30	0.46	weak*/fair wind	0.25	0.43
serious*/severe pain	0.28	0.45	light*/slight smile	0.23	0.42
main*/major storm	0.28	0.45	thick*/heavy scent	0.23	0.42
real*/true colors	0.25	0.43	tough*/heavy seas	0.20	0.40
clear/clean* skin	0.25	0.43	rare*/slim chance	0.17	0.37
clear*/clean election	0.24	0.43	dirty*/ugly truth	0.15	0.36
Total Mean	**0.46**	**0.47**	**Total Mean**	**0.36**	**0.46**

Another problematic IAN collocation for the learners was *ugly truth* which was guessed correctly by only 15.0% of the learners. A probable reason for this is that the learners may have recalled the L1 *deoleowoon jinshil* (i.e., 더러운 진실) which is more closely equivalent to **dirty truth* when translated to L2. Another possibility for the low mean response can be attributed to the way Korean L2 high school learners of English tend to associate *ugly* as a word to describe someone's unpleasant facial features. As such, the example demonstrates that the learners may lack knowledge of polysemous adjectives (e.g., *ugly*) whose use may become restricted by the use of particular nouns (e.g., *truth*). For this reason, the results indicate that learners may need to be instructed sometimes to look beyond first entry dictionary meanings, and try to establish a firm association between a particular adjective and a particular noun.

A different kind of explanation needs to be provided for the VN collocations, since the results are different from those of AN. The results also contradict those claimed in previous research (Kim, 2012; Nesselhauf, 2003; Wolter & Gyllstad, 2011; Yamashita & Jiang, 2010) where the incongruent collocations were found to be more difficult for L2 learners to acquire due to the potential interference from L1. When the mean responses of the VN items were examined (See Table 4), the two lowest mean values recorded for CVN were **recover/restore faith* (M = 0.15) and *obtain/*gain data* (M = 0.26). While the Korean equivalent of *restore faith* is *silloereur hoebokhada* (i.e., 신뢰를 회복하다), a plausible explanation for the confusion is that the learners had not been able to distinguish the different nuances of meaning for the synonymous collocates (i.e., *recover* vs. *restore*), both of which can be translated as 'recover' (i.e., 회복하다) in Korean. For *obtain/*gain data*, most learners (73.7%) responded by marking *gain* rather than *obtain* likely because the given collocates may both may mean 'get' or *eodda* (얻다) in Korean. Other examples of synonymous collocates are *extend* and *expand* for *expand the market*. As such, the results demonstrate that the synonymous collocates (e.g., *recover, restore*) would have prompted competition between the words in the mental lexicon as the learners were trying to find the best matching collocate for a node. The results also demonstrate that L2 CAN collocations, in spite of their congruency with L1, can be sources of problems. These were cases when the learners often thought they knew a word, but it was in fact a word they did not know properly (Laufer, 1997).

TABLE 4
Mean Responses for 'Verb + Noun' Collocations

Congruent Collocations			Incongruent Collocations		
Verb+Noun	*M*	*SD*	Verb+Noun	*M*	*SD*
exam*/test my patience	0.68	0.47	open*/have a party	0.73	0.45
stay*/keep a balance	0.62	0.49	see*/take an exam	0.69	0.46
erase*/remove the obstacle	0.59	0.49	eat*/take your medicine	0.68	0.47
achieve/fulfill* success	0.54	0.50	sleep*/take a nap	0.60	0.49
raise*/increase speed	0.47	0.50	collect*/make a fortune	0.54	0.50
erase*/remove dust	0.47	0.50	talk*/have a chat	0.52	0.50
take*/accept the truth	0.46	0.50	give*/make a change	0.51	0.50
grow/raise* crops	0.43	0.50	do*/have a debate	0.49	0.50
raise/increase* your voice	0.43	0.50	do*/make an effort	0.47	0.50
ended/finished* their relationship	0.41	0.49	see*/reach an agreement	0.46	0.50
mend*/fix the problem	0.41	0.49	keep/have* the change	0.42	0.50
lifted*/raised expectations	0.41	0.49	catch*/take a chance	0.41	0.49
erase/remove* the memory	0.41	0.49	stop*/hold your breath	0.40	0.49
start/begin* the engine	0.35	0.48	give*/cause damage	0.39	0.49
provide/supply* protection	0.32	0.47	earn*/get enough allowance	0.38	0.49
got*/gained popularity	0.32	0.47	take*/get surgery	0.36	0.48
watching*/seeing a patient	0.32	0.47	clean*/make the bed	0.28	0.45
extend*/expand our market	0.29	0.46	keep*/meet that deadline	0.25	0.43
obtain/gain* data	0.26	0.44	brought*/caused a sensation	0.25	0.43
recover*/restore faith	0.15	0.36	know*/see your point	0.19	0.39
Total Mean	**0.42**	**0.48**	**Total Mean**	**0.45**	**0.48**

Another explanation for the higher mean values for the IVN than the CVN collocations can be attributed to the responses of the high proficiency learners. For the IVN *see*/take an exam* (*M* = 0.69), IVN *eat*/take your medicine* (*M* = .68) and IVN

sleep/take a nap* (M = 0.60), the high proficiency learners scored full marks (i.e., 1.00) (See Table 4). Post-hoc multiple comparisons indicated differences between the high-mid proficiency group learners for *take an exam* (p = .027) and *take a nap* (p = .010) at a significant level (p < .05), and at a borderline significant level for *take your medicine* (p = .050). These three collocations also illustrated significant differences between the high-low proficiency groups (p < .05). Although the items were incongruent and were thus expected to be relatively more challenging for the learners, the collocations that took 'take' as the node were evidently more familiar to the high proficiency learners. This adds to the explanation of how higher mean scores were produced for the IVN collocations than those of the CVN collocations. Another possibility is that the learners knew the specific nouns (i.e., *exam*, *medicine*, and *nap*) well enough to see that they would prompt the verb 'take.' All in all, the pattern of the responses for the 'take + NP' collocations (i.e., M = 0.60 ~ 0.69) indicates that they were well known to the learners, and we can attribute this to how common they are in the nationally authorized English high school textbooks. For instance, the researchers were able to find a variety of expressions in the most recent English textbooks (i.e., *English 1*: 6 publishers and *English 2*: 6 publishers) where 'take' was often used as the verb of a collocation (e.g., *take a photo, take pictures, take a walk, take time, take turns, take action, take advice*, and *take place*). This presumably had an influence on the level of familiarity the learners exhibited towards collocations that took the form of 'take + NP.' Similarly, when Lee (2010) analyzed the corpus of English textbooks used in Korean high schools, she found 'take + NP' to take up 50% of the lexical chunks that included 'take' as the node. She also found 'take' to collocate with words, such as 'medicine' or 'vitamin D,' and 'test' or 'exam.'

Other than the frequency of the 'take + NP' items, it seems that the learners may have used guessing strategies to comprehend the meaning of 'take + NP' items. Learners may have paid attention to the noun rather than the verb since most of the meaning is found in the noun, and not in the verb. This is likely to occur when delexical verbs (e.g., *have, take, make, give*) constitute collocations.

2. Knowledge of Collocations and L2 Proficiency

To deal with RQ 2, a part of the study consisted of researching the relationship between learners' knowledge of collocations and L2 proficiency. With information on learner proficiency by stanine levels of the learners' mock CSAT, the learners were categorized into three groups (See previous Table 1). Examining mean differences between groups with one-way ANOVA produced significant results ($F(2, 130) = 29.96$, p = .000) (Table 5). Post-hoc multiple comparisons consistently showed that the higher

proficiency learners had obtained higher scores on the collocation test, and the differences were significant between all three proficiency groups (i.e., high, mid, low) at the level of $p < .01$. This coincides with the results of previous studies that general L2 proficiency is a contributing factor for obtaining knowledge of L2 collocations (Al-Zahrani, 1998; Kim, 2012; Yamashita & Jiang, 2010). The results warranted further analysis for the different types of collocations.

TABLE 5
Knowledge of Collocations by Proficiency Groups

	N	Mean	*SD*	95% Confidence Interval for Mean		*F*	*Sig.*
				Lower Bound	Upper Bound		
1	17	0.62	0.11	0.03	0.56	29.96	.000
2	86	0.43	0.14	0.02	0.40		
3	30	0.29	0.15	0.03	0.24		
Total	**133**	**0.42**	**0.17**	**0.01**	**0.39**		

Note: *$p < .05$; 1 = High, 2 = Mid, 3 = Low

Respectively analyzing congruent and incongruent collocations indicated that the learners' knowledge of collocations was different according to L2 proficiency (Congruent: $F(2, 130) = 25.865$, $p = .000$; Incongruent: $F(2, 130) = 27.258$, $p = .000$) (Table 6). Further post-hoc multiple comparisons for both congruent and incongruent collocations indicated that there were significant differences between all three proficiency groups ($p < .01$). The results suggest that L2 general language proficiency is a strong factor in the acquisition of collocations. The next logical step of the analysis was also to examine learners' knowledge by the different parts of speech (i.e., AN and VN).

The three proficiency groups all demonstrated different levels of collocation knowledge for CAN ($F(2, 130) = 14.373$, $p = .000$), IAN ($F(2, 130) = 14.996$, $p = .000$), CVN ($F(2, 130) = 25.020$, $p = .000$) and IVN ($F(2, 130) = 28.063$, $p = .000$) (Table 7 & 8). Post-hoc multiple comparisons for CAN and IAN all indicated that there were significant differences between the three proficiency groups ($p < .01$). Post-hoc multiple comparisons for both CVN and IVN also indicated significant differences between the three proficiency groups ($p < .01$).

TABLE 6
Knowledge of Collocations by Congruency

		Mean	*SD*	95% Confidence Interval		*F*	*Sig.*
				Lower Bound	Upper Bound		
Congruent	1	0.64	0.11	0.58	0.70	25.865	.000
	2	0.44	0.16	0.41	0.48		
	3	0.31	0.16	0.25	0.37		
	Total	**0.44**	**0.18**	**0.41**	**0.47**		
Incongruent	1	0.61	0.13	0.54	0.68	27.258	.000
	2	0.41	0.15	0.38	0.45		
	3	0.28	0.16	0.22	0.34		
	Total	**0.41**	**0.18**	**0.38**	**0.44**		

Note: * $p < .05$; 1 = High, 2 = Mid, 3 = Low

TABLE 7
Knowledge of Collocations by Congruency for Adj +Noun Collocations

		Mean	*SD*	95% Confidence Interval		*F*	*Sig.*
				Lower Bound	Upper Bound		
Congruent Adjective+ Noun	1	0.61	0.11	0.55	0.67	14.374	.000
	2	0.46	0.15	0.43	0.49		
	3	0.36	0.18	0.29	0.43		
	Total	**0.46**	**0.17**	**0.43**	**0.49**		
Incongruent Adjective+ Noun	1	0.53	0.15	0.45	0.61	14.996	.000
	2	0.37	0.16	0.33	0.40		
	3	0.27	0.17	0.20	0.33		
	Total	**0.36**	**0.17**	**0.33**	**0.39**		

Note: ** $p < .01$; 1 = High, 2 = Mid, 3 = Low

TABLE 8
Knowledge of Collocations by Congruency for Verb + Noun Collocations

		Mean	*SD*	95% Confidence Interval		*F*	*Sig.*
				Lower Bound	Upper Bound		
Congruent Verb+Noun	1	0.67	0.14	0.60	0.74	25.020	.000
	2	0.42	0.20	0.38	0.47		
	3	0.26	0.19	0.19	0.33		
	Total	**0.42**	**0.22**	**0.38**	**0.46**		
Incongruent Verb+Noun	1	0.69	0.15	0.62	0.77	28.063	.000
	2	0.46	0.19	0.42	0.50		
	3	0.29	0.18	0.22	0.35		
	Total	**0.45**	**0.22**	**0.41**	**0.49**		

Note: ** $p < .01$; 1 = High, 2 = Mid, 3 = Low

On the whole, the analysis indicates that L2 general proficiency is a variable that has a strong relationship to learning L2 collocations. That is, regardless of congruency, there were valid differences in the learners' collocational knowledge according to general L2 proficiency. The pattern of results suggests that collocation knowledge is likely to develop on a continuum with increasing general proficiency. In other words, the results also suggest that even EFL learners, who often lack sufficient input to the target language, can develop accurate and fluent collocational knowledge, when they aim to develop general L2 proficiency. Nevertheless, when teaching, Boers et al. (2014) point out that a prerequisite for learning collocations is to have them brought to the learners' attention when they are presented either explicitly or incidentally. This is because collocations (e.g., *do the dishes*) are often composed of common words (e.g., high-frequency lexical items), and thus not easily noticeable by L2 learners (Nesselhauf, 2003).

An implication that can be gained from the results is that familiarity with collocations cannot simply be acquired through rote learning of collocation items or a list, but that general language ability needs to become the foundation for learning the more complex types of collocations. However, for the collocations to be acquired, learners will need numerous exposures to the collocation items, whether through message-focused activities or deliberate learning of pre-selected collocations. For instance, multiple encounters with collocations are likely to establish a firm association between a

particular verb and a particular noun. However, Boers et al. (2014) raises the point that "the chances of a learner meeting the same collocation often enough in a relatively short span of time are pretty slim" (p. 56). For these reasons, the deliberate learning of pre-selected collocations (Laufer & Girsai, 2008) may have to accompany the practice of extensive reading and listening for learning collocations.

V. CONCLUSION

The results proved to show that the incongruent collocations can be considered more difficult for the learners to retrieve than the congruent collocations. For the incongruent collocations, it was apparent that the gap that exists between the constituent words of L1 and L2 collocations may lead to the erroneous use of L2 collocations. On the other hand, for the congruent collocations, it was the presentation of synonymous collocates (e.g., *expand* and *extend*) that were sometimes the source of confusion and error. The learners are expected to have contemplated the optional collocates before being able to retrieve the most appropriate collocation. In particular, as with the CAN collocations, they became the source of the problem when the meanings of the adjectives were considered only in the way the learners understood them (e.g., first meaning in a dictionary entry), and the learners were unaware of how a node word (e.g., surface) could impose a restriction on the adjective (e.g., **tough/rough surface*). This also implies that learners may need to be trained to acquire L2 collocations in ways that do not always require mediation through L1, since learners had difficulty retrieving the collocations whether they were dealing with congruent or incongruent collocations (Williams & Cheung, 2011; Wolter 2006). Meanwhile, the relationship between the RCT mean values and the learners' general proficiency also indicate that obtaining English proficiency is an important prerequisite for building lexical knowledge of collocations.

Some implications can be drawn from the study. The study points out that the teaching of collocations would need to be done in a more systematic way. One of the important reasons is that the amount of L2 input EFL learners receive is often limited to the classroom, unlike those in ESL (English as a Second Language) contexts, whose knowledge of collocations can be developed more naturally. For EFL learners, when collocations are met in an arbitrary manner, the learners may never be able to recognize the "legitimacy of an incongruent collocation over a congruent one" (Yamashita & Jiang, 2010, p. 652). For teachers would need to take into account that it is more than one factor that is likely to influence the learning burden of collocations (Peters, 2016; Webb & Kagimoto, 2011). First, learners need to be instructed that relying on L1 equivalents is detrimental to forming knowledge of incongruent collocations, and that

this problem can be exacerbated by their tendency to treat L2 collocations mainly as the combination of single word items. Second, learners' attention needs to be drawn to the polysemy of adjectives/verbs in the collocations and their relations to L1. As we saw previously, learners will mistakenly retrieve the prototypical meaning of a collocate and when the collocation is not stored holistically, learners are likely misinterpret the collocation. Teachers should thus highlight to learners a word's polysemy and show that adjectives and verbs in combination with nouns might have a different meaning.

Although the researchers tried to counterbalance the exposure rate of the target collocations between those that are congruent and incongruent by accessing frequency information from COCA, it seemed that the learners were already familiar with specific types of collocations that had in fact appeared frequently in the recently published high school textbooks. Particular items were already known to some of the learners regardless of congruency, contributing to the higher mean responses for the IVN collocations. However, the different exposure rates that emerged may not be surprising after all since the frequency information accessed at COCA is ultimately a representation of native-speaker corpora. At the time of study, COCA was considered the most valid source of frequency information for designing RCT, and selecting collocations. For a reliable development of a collocation test and selection of items to be used with a specific group of learners, consideration hereafter should be further made for how the learners are exposed to other educational materials within the EFL context by which some collocation items can be ruled out for testing.

Some possible limitations of the study exist. The study was restricted to testing learners on a number of AN and VN collocations. Future studies may need to test learners on a larger variety of collocations that may have general validity so that the results may be applicable to a larger EFL learner population. Also, although effort was made to obtain information on the learners' L2 proficiency, their L1 proficiency may have influenced the way they translated the L2 collocations since learners may have retrieved different forms of multiword items in L1. This unexplored variable may need to be examined or controlled when conducting a study of a similar kind, since having to retrieve an L2 collocation may render an L1 translation (Yamashita & Jiang, 2010).

REFERENCES

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.

Al-Zahrani, M. S. (1998). *Knowledge of English lexical collocations among male Saudi college students majoring in English at a Saudi university*. Unpublished doctoral

dissertation, Indiana University of Pennsylvania, Pennsylvania.

Benson, M. (1989). The structure of the collocational dictionary. *International Journal of Lexicography*, *2*(1), 1-14.

Boers, F., Eyckmans, J., Kappel, J., Stengers, H., & Demecheleer, M. (2006). Formulaic sequences and perceived oral proficiency: Putting a lexical approach to the test. *Language Teaching Research, 10*(3), 245–61.

Boers, F., Lindstromberg, S., & Eyckmans, J. (2014). Some explanations for the slow acquisition of L2 collocations. *Vigo International Journal of Applied Linguistics, 11*, 41-62.

Chen, P. C. (2002). *A corpus-based study of the collocational errors in the writings of the EFL learners in Taiwan*. Unpublished master's thesis, National Taiwan Normal University, Taipei.

Chon, Y. V., & Lee, S. (2015). L2 collocational knowledge by congruency and writing proficiency of EFL university learners. *English Language & Literature Teaching*, *21*(1), 49-74.

Conklin, K., & Schmitt, N. (2008). Formulaic sequences: Are they processed more quickly than nonformulaic language by native and nonnative speakers? *Applied Linguistics, 29*(1), 72–89.

Conzett, J. (2000). Integrating collocation into a reading and writing course. In M. Lewis (Ed.), *Teaching collocation: Further developments in the lexical approach* (pp. 70-87). Hove: Language Teaching.

Davies, M. (2008-2012). The Corpus of Contemporary American English: 450 million words, 1990-present. Retrieved May 6, 2013 from the World Wide Web: http://corpus.byu.edu/coca/.

Doosan Dong-A prime English-Korean dictionary (6th ed.). (2011). Seoul: Doosan Dong-A.

Erman, B., & Warren, B. (2000). The idiom principle and the open choice principle. *Text-Interdisciplinary Journal for the Study of Discourse, 20*(1), 29-62.

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

Howarth, P. (1998). Phraseology and second language proficiency. *Applied Linguistics, 19*(1), 24-44.

Jiang, N. (2004). Semantic transfer and its implications for vocabulary teaching in a second language. *The Modern Language Journal, 88*, 416-432.

Jung, Y.–C., Park, M.–S., & Kim, E.–I. (2010). A study of Korean high school students' collocation knowledge of English synonymous adjectives. *Journal of Language Sciences, 17*(3), 139-161.

Jung, Y.–C., Sur, R.–I., & Kim, E.–I. (2007). A corpus-based study of Korean college students' collocation ability of English synonymous verbs. *The New Korean Journal of English Language & Literature, 49*(2), 175-194.

Kim, K.–M. (2012). L1 influence on L2 collocational knowledge: By level of receptive vocabulary knowledge and general language proficiency. *Korean Journal of Applied Linguistics, 28*(3), 1-38.

Kim, N.–B. (2003). An investigation into the collocational competence of Korean high school EFL learners. *English Teaching, 58*(4), 225-248.

Kim, N.–B. (2005). *A study on collocation-based English vocabulary teaching and learning*, Unpublished doctoral dissertation, Chungnam National University, Daejeon, Korea.

Kim, N.–B. (2010). The effects of the explicit lexical collocation instruction on the grammatical collocation knowledge: Focusing on preposition collocations. *English Language & Literature Teaching, 16*(3), 141-160.

Ko, S. (2010). The e*ffect of collocation-based vocabulary instruction on the writing ability of middle school students*. Unpublished master's thesis, Hanyang University, Ansan, Korea.

Kroll, J. F., & Stewart, E. (1994). Category interference in translation and picture naming: Evidence for asymmetric connections between bilingual memory representations. *Journal of Memory and Language, 33*, 149–174.

Laufer, B. (1997). The lexical plight in second language reading: Words you don't know, words you think you know and words you can't guess. In J. Coady & T. Huckin (Eds.), *Second language vocabulary acquisition: A rationale for pedagogy* (pp. 20-34). Cambridge: Cambridge University Press.

Laufer, B., & Girsai, N. (2008). Form-focused instruction in second language vocabulary learning: A case for contrastive analysis and translation. *Applied Linguistics*, *29*(4), 694-716.

Lee, J.–W., & Kim, N.–B. (2005). The effects of the collocation-based vocabulary instruction on English vocabulary development. *English Education, 60*(2), 111-135.

Lee, S. R. (2010). *A corpus based analysis of high-frequency verbs collocations in high school textbooks*. Unpublished master's thesis, Yonsei University, Seoul.

Lewis, M. (Ed.). (2000). *Teaching collocation: Further developments in the lexical approach*. London: Language Teaching.

Liu, C. P. (1999). An analysis of collocational errors in EFL writings. In *The Proceedings of the Eighth International Symposium on English Teaching* (pp. 483-494). Taipei: Crane.

Liu, D. (2010). Going beyond patterns: Involving cognitive analysis in the learning of collocations. *TESOL Quarterly*, *44*(1), 4-30.

Martinez, R., & Murphy, V. (2011). Effect of frequency and idiomaticity on second language reading comprehension. *TESOL Quarterly, 45*(2), 267–290.

McCarthy, M., & O'Dell, F. (2005). *English collocations in use*. Cambridge: Cambridge University Press.

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.

Nizonkiza, D. (2011). The relationship between lexical competence, collocational competence, and second language proficiency. *English Text Construction, 4*(1), 113-145.

Oxford collocations dictionary for students of English. (2010). Oxford: Oxford University Press.

Oxford advanced learners' English-Korean dictionary. (2008). Oxford: Oxford University Press.

Peters, E. (2016). The learning burden of collocations: The role of interlexical and intralexical factors. *Language Teaching Research, 20*(1), 113-138.

Webb, S., & Kagimoto, E. (2009). The effects of vocabulary learning on collocation and meaning. *TESOL Quarterly, 43*, 55-77.

Webb, S., Newton, J., & Chang, A. (2013). Incidental learning of collocation. *Language Learning, 63*(1), 91-120.

Williams, J. N., & Cheung, A. (2011). Using priming to explore early word learning. In P. Trofimovich & K. McDonough (Eds.), *Applying priming methods to L2 learning, teaching and research: Insights from psycholinguistics* (pp. 73-103). Amsterdam: John Benjamins.

Wolter, B. (2006). Lexical network structures and L2 vocabulary acquisition: The role of L1 lexical/conceptual knowledge. *Applied Linguistics, 27*(4), 741-747.

Wolter, B., & Gyllstad, H. (2011). Collocational links in the L2 mental lexicon and the influence of L1 intralexical knowledge. *Applied Linguistics, 32*(4), 430-449.

Wray, A. (2008). *Formulaic language: Pushing the Boundaries*. Oxford: Oxford University Press.

Yamashita, J., & Jiang, Y. (2010). L1 Influence on the acquisition of L2 collocations: Japanese ESL users and EFL learners acquiring English collocations. *TESOL Quarterly, 44*(4), 647-668.

Examples in: English
Applicable Language: English
Applicable Level: Secondary

Eun-Sun Kim
Dept. of English Education
College of Education
Hanyang University
222 Wangshimli-Ro
Seongdong-Gu
Seoul 04763
Email: mornmoon@naver.com

Yuah V.Chon
Dept. of English Education
College of Education
Hanyang University
222 Wangshimli-Ro
Seongdong-Gu
Seoul 04763
Email: vylee52@hanyang.ac.kr

Received in January 15th, 2017
Reviewed in February 26th, 2017
Revised version received in March 5th, 2017