

## Structural Use of Lexical Bundles in the Rhetorical Moves of L1 and L2 Academic Writing

Ji-yoon Hong

(Hankuk University of Foreign Studies)

**Hong, Ji-yoon. (2019). Structural use of lexical bundles in the rhetorical moves of L1 and L2 academic writing. *English Teaching*, 74(3), 29-54.**

This corpus-driven longitudinal study investigates the structural use of lexical bundles in published research article (RA) introductions in applied linguistics written by English experts and Korean graduate students across two different levels of study. Frequency-based bundles were retrieved from a corpus of 200 published RA introductions and two corpora of 46 and 49 introductions of term papers written at two time points of the first and fourth semester of graduate course. In a further step, the structures of the bundles in different rhetorical moves of RA introductions were analyzed to reveal the developmental patterns in bundle use. The analyses show that the Korean graduate students are in the developmental process of academic writing featured by a shift from clausal style to phrasal style as their academic level advances. The results also suggest that the students have difficulty in appropriate bundle use in specific rhetorical moves even at the later academic level of graduate coursework. The pedagogical implications of L2 students' developmental order are discussed.

**Key words:** lexical bundle, rhetorical move, academic writing, learner corpus, interlanguage development

### 1. INTRODUCTION

The challenge of learning to write academic texts in English and writing publishable research articles (RAs) is becoming an increasingly important issue for novice researchers including second language (L2) graduate students in the ever-competitive world of academic publishing (Hyland, 2015). For this reason, genre analysis has been widely used in English for Academic Purposes (EAP) research. Particularly, many corpus-based studies over the past two decades have conducted analysis of moves, which are rhetorical structures with specific communicative functions, for RAs using Swales's (1990, 2004)

move framework (e.g., Biber, Connor, & Upton, 2007; Kanoksilapatham, 2007). Several researchers have been also focusing on identifying a list of words or multiword expressions that can characterize the different rhetorical moves of RA (Kanoksilapatham, 2003; Yang & Allison, 2003). Among multiword expressions, lexical bundles refer to sequences of word forms that commonly go together in natural discourse such as *as a result of* or *the way in which* (Biber, Johansson, Leech, Conrad, & Finegan, 1999). Lexical bundles have been widely investigated in that they play an important role in reflecting the difference between genres or disciplines (e.g., Biber, Conrad, & Cortes, 2004; Durrant, 2017; Grabowski, 2015; Hyland, 2008b). Identifying lexical bundles in the rhetorical moves of RA would empirically demonstrate which expressions are frequently employed to initiate rhetorical moves in RA. In addition, since moves as main building blocks of a genre (Biber et al., 2007) and lexical bundles as text building blocks of coherent discourse (Biber, Conrad, & Leech, 2002) have similar characteristics, a description of the relationship between them in a particular genre is undoubtedly needed in order to teach a genre to EAP students in effective and holistic ways (Cortes, 2013).

Despite the abundance of studies on move analysis or lexical bundles in RA, there is still little research empirically investigating the bundle use in the specific rhetorical moves (e.g., Cortes, 2013; Mizumoto, Hamatani, & Imao, 2017) or L2 learner development in this bundle use in the rhetorical moves in RA in terms of structure across different levels of studies compared to first language (L1) experts. Therefore, the current study aims to fill the gap by investigating the structural use of lexical bundles in the rhetorical moves in introductions of published RAs produced by English experts and those of term papers produced by Korean EAP students at two different levels of graduate course. The following research questions guide this study:

1. What are the structural forms of the bundles of four+ words identified in the introductions of published RAs and those of two groups of student term papers?
2. How do the introductions of published RAs and those of two groups of student term papers differ in the structural use of the lexical bundles in each move?

By examining whether the Korean EAP students' structural use of lexical bundles in different rhetorical moves of RA introductions approximate to expert writers' usage in line with increase of their levels of study, the current corpus-driven study contributes to providing greater insight into genre-based writing instruction considering EAP students' developmental patterns for structural bundle use.

## 2. REVIEW OF THE LITERATURE

### 2.1. Language Development in SLA and Lexical Bundles

It has generally observed that advanced and fluent writing is characterized by appropriate and frequent use of formulaic language which maintains language users' identity in an academic community (Wray, 2002). Among several terms of formulaic language, lexical bundles have been referred to as the most frequently recurring sequences of words in a given genre retrieved by corpus-driven approach with specified frequency and distribution criteria (Biber et al., 1999). The words in these lexical bundles occur together more often than would be expected by chance, and they may or may not consist of a complete syntactic unit such as *is likely to* in 'our response to the first question is that the nature of teacher research engagement *is likely to* vary globally depending on the specific conditions and factors' or *the fact that the* in 'despite *the fact that the* aforementioned definitions of fluency may apply to both native and nonnative speech, fluency assessment has thus far mostly been aimed at nonnative speakers'.

As Coxhead and Byrd (2007) pointed out, a writer who can successfully comprehend and produce texts using lexical bundles correctly is likely to be considered as a fluent reader and writer of the language. However, the multi-word complexity of the bundles which combine several different parts of speech makes new academic writers challenged to master these bundles. Research has found that the inability to use bundles appropriately deprives students and even scholars of a valuable text construction resource and to some extent excludes them from the academic discourse community to which they seek to belong (Cortes, 2004; Salazar, 2014; Staples, Egbert, Biber, & McClair, 2013).

In the field of EAP and L2 writing, recognizing that the unnatural of papers written by L2 students is due to a lack or misuse of lexical bundles, researchers have shown great interest in understanding how the lexical bundles used by L2 writers and native English-speaking writers differ (Ädel & Erman, 2012; Chen & Baker, 2010, 2016; Staples et al., 2013; Wei & Lei, 2011). Staples et al. (2013) found that as learners gain proficiency in a L2 they tend to employ fewer lexical bundles by examining bundle variation across three proficiency levels in prompted TOEFL writing. They found that the lowest level students used more bundles overall containing considerable prompt bundles while the intermediate level students used more non-prompt bundles. Compared to the highest-level students, the intermediate level students used significantly more bundles overall, and this difference remained once the prompt bundles were removed. These findings support Second Language Acquisition (SLA) theory that developmental sequences begin with repeating unanalyzed language such as memorization and one-to-one form-function mapping and move toward self-constructed language as learners gain proficiency in a L2 (Ellis, 2002,

2006; Myles, Hooper, & Mitchell, 1998).

Drawing on a corpus of published academic texts, a corpus of L1 student academic writing and a corpus of L2 Chinese student academic writing, Chen and Baker (2010) analyzed the structures and functions of lexical bundles in L1 and L2 academic writing. One of the interesting results they found was the relationship between the number of lexical bundles and writing proficiency. The number of lexical bundles increased with advancing writing proficiency, which was the case both for the types and tokens of lexical bundle used. This result was contrary to some of the results reported in the literature (De Cock, 2000; Hyland, 2008a), possibly due to removal of overlapping bundles or context-dependent ones such as topic-related clusters (e.g., *in the Hong Kong*), relatively smaller corpus size or less strict dispersion requirement. As for structural comparison, it has been found that more VP-based bundles were contained in L1 and L2 student writing than in L1 expert writing, which appears to be a sign of immature writing, while more NP-based bundles were used in L1 expert writing than in L1 and L2 student writing. The divergent use of lexical bundles by L2 students from that of professional writers was also found in Wei and Lei's (2011) research, in which it was argued that the overuse of passive structures and the underuse of anticipatory *it* structures of bundles than the professional writers may be due to the L2 students' preference for the impersonality in their academic prose. Ädel and Erman (2012) also confirmed a general pattern found by prior bundle research, which is that non-native speakers show a more restricted repertoire of lexical bundles than native speakers. Particularly, as for structural use of bundles, more varied means of expression among the native speakers were found in unattended *this* constructions, existential *there* constructions, passive constructions, negated patterns (e.g., *not be able to*, *there is no evidence*) and *fact*-headed bundles (e.g., *the fact that they*) compared to non-native speakers.

The use of lexical bundles across various proficiency levels from L1 Chinese learners was investigated in Chen and Baker's (2016) research, in which proficiency was determined by a robust rating procedure that is often used in high-stakes tests instead of the traditional approach of utilizing extra-linguistic judgment. They found that lower-level writing shares more features with conversation due to bundle use which is more verb-heavy and personally involved with more colloquial quantifiers while more proficient writing shows an opposite pattern, having a more impersonal tone with greater use of nominal components in lexical bundles and also sharing more academic or literate bundles with the register of academic prose.

Although these cross-sectional studies so far have suggested a developmental sequence for some aspects of bundle use by L2 writers, there is still little longitudinal research exploring L2 learner development by analyzing writings produced by same L2 students at different levels of studies compared to English L1 expert writers.

## 2.2. Move Analysis in Academic Writing

Genres are considered abstract, socially recognized communicative strategies rather than the culture at large which represent how writers typically handle recurrent situations (Hyland, 2007, 2016). Genres are dynamic flexible, and open to change in response to community members' needs and contextual changes. However, through repeated exposure to these genres, members of a specific genre community acquire the prototypical schematic structures and linguistic features used to realize them which are required for their specific purposes. As a powerful tool in text analysis that provides insights into these important characteristics of genres within a specific discipline, genre analysis has been a critically influential topic of much research in EAP and L2 writing over the last few decades (Cotos, 2014; Hyland, 2013; Paltridge, 2013).

Move analysis, a specific type of genre analysis, first proposed by Swales (1981, 1990), is exemplified through the *Create A Research Space* (CARS) model, which is an analytic framework for describing RA introductions. A move as a functional unit in a text which achieves the overall communicative purpose of the genre when working together is realized by as short as a clause and as long as a paragraph (Connor, Davis, & De Rycker, 1995). The general organization of a text can be described as consisting of several moves, with some conventional/obligatory moves occurring more often than other moves which are optional. A move contains one or more specific steps which realize the move when combined (Swales, 1990). In Swales's (2004) revised CARS model for RA introductions, three move structures were introduced for more comprehensive analysis and identification of a higher number of steps: Move 1, *Establishing a territory*; Move 2, *Establishing a niche*; and Move 3, *Presenting the present work*. Swales's initial motivation for developing this analytical framework was to help L2 students for whom English is not their first language to improve their reading and writing in English and compensate for their perceived constraints inherent in the academic discourse. Many EAP researchers have thus applied this scheme in order to identify the linguistic features characterizing the underlying generic structure of RA sections (e.g., Cortes, 2013; Cotos, Haufman, & Link, 2017; Durrant & Mathews-Aydinli, 2011; Henry & Rosenberry, 2001; Kanoksilapatham, 2005, 2007; Le & Harrington, 2015; Lim, 2006; Swales, 1981). These research attempts to fill the function-form gap involve establishing the most salient linguistic items or patterns, occurring in a specific rhetorical context in a RA, through which a given communicative function can be interpreted in a highly predictable manner for readers (Moreno & Swales, 2018).

### 2.3. Lexical Bundles in Rhetorical Moves

The identification of lexical items and lexico-grammatical patterns used to signal the onset of rhetorical moves has gained considerable attention for researchers studying discourse organization. Swales (1981) identified some lexical items as key markers for moves by stating that in the first move of introductions, authors use expressions which denote interest or importance for the communicative function of asserting centrality.

Recently, by adopting corpus-driven approach, several empirical studies for identifying lexical bundles in the rhetorical moves of RA introductions have demonstrated which expressions are frequently used to initiate rhetorical moves in this section of RAs. Among these studies, Cortes (2013) addressed a noticeable concept which brings lexical bundles and rhetorical moves together. That is, both linguistic features have been considered building blocks for constructing discourse. Lexical bundles have been defined as important text building blocks of discourse in spoken and written register which represent characteristic features of different registers and genres (Biber & Barbieri, 2007; Biber et al., 2002; Conrad & Biber, 2004; Hyland, 2008b). Biber et al. (2007) also pointed out that lexical bundles are seen as lexico-grammatical building blocks associated with basic functions used to combine the text together. In a similar vein, rhetorical moves have also been considered main building blocks of a genre which can be used to teach novice writers how to produce successful texts in that specific genre (Biber et al., 2007; Dudley-Evans, 1995).

Despite recent several studies in which a similar approach has been suggested and conducted considering their similar characteristics as building blocks (Durrant, 2017; Durrant & Mathews-Aydnli, 2011; Flowerdew, 2009; Le & Harrington, 2015), Cortes' (2013) research was the first to explicitly integrate rhetorical moves with lexical bundles in RAs. She found a strong connection between the communicative functions of certain bundles and the rhetorical purposes of the moves and steps in which they were found. Mizumoto (2016) extended the analysis of the location and specific function of bundles to abstracts and all four IMRD sections of RAs by finding that a given bundle may be tied to individual moves in more than one RA section and achieve a range of rhetorical purposes. More recently, Mizumoto, Hamatani, and Imao (2017) explored the pedagogical implications of the correspondence between moves and bundles by developing an online writing support tool which allows the user to search for the most frequently used bundles in different moves of RA. More empirical evidence of the link between rhetorical moves and their linguistic realizations across different disciplines was found in Omidian, Shahriari, and Siyanova-Chanturia's (2018) research, in which the members of different academic domains have different priorities for representing their research in RA abstracts.

These studies so far have provided evidence for the functional use of lexical bundles in

the specific rhetorical moves to build discourse by different academic communities and argued against identification of moves and their linguistic features solely based on qualitative evidence drawn from subjective observations. However, few studies to date have little attention to how EAP students at different levels of study show the structural use of lexical bundles in the rhetorical moves different from English expert writers from a L2 developmental perspective. Specifically, more empirical research is needed to develop a better understanding of the structural use of lexical bundles in the rhetorical moves by analyzing longitudinal learner corpora compared to English expert corpora.

### **3. METHODOLOGY**

#### **3.1. Data Collection**

For comparable or equivalent corpora of this study, similar contextual conditions of the corpora such as field of study, text form, genre, mode, time-frame for collection, participants to the highest degree rather than exactly the same were satisfied (Moreno, 2008). Thus, introductions were selected for analysis among several different sections of RA since they are known to elucidate overall organization and a purpose of RA clearly (Swales, 1990). In addition, considering the time-frame for collection and the fact that genres evolve over time, the three corpora were compiled using the introductions of RAs and term papers published or written from 2011 to 2016 (Guinda, 2015).

Therefore, for the analysis, a reference corpus against two groups of L2 student corpora (hereafter English L1 corpus) was composed of 200 published empirical RA introductions selected from six prestigious journals in applied linguistics with high impact factors above 1.0, *System*, *TESOL Quarterly*, *Language Learning*, *English for Specific Purposes*, *Applied Linguistics*, and *Journal of Second Language Writing*, whose writers were regarded experts. The two groups of L2 student corpora (hereafter English L2-1 corpus and English L2-2 corpus) consisted of 46 and 49 introductions of term papers written by same groups of native Korean graduate students at their first and fourth semester, respectively. In the face-to-face interview for obtaining information about their previous L2 writing experiences, the students reported that they were all majoring in English education as MA students at several graduate schools in Korea and had to submit term papers as assignments during semester to prepare for a thesis. Moreover, they had no experience of publish RAs in English and their L2 English proficiency levels corresponded to intermediate to advanced according to the Common European Framework of Reference for Language (CEFR), one of the most influential English profile frameworks (Council of Europe, 2001). That is, the scores of most students ranged 87-109 on the TOEFL iBT and 785-990 on the

TOEIC listening and reading tests total and those of some students ranged 110-120 on the TOEFL iBT and 490-495 on the TOEIC listening tests, which corresponded to the intermediate and advanced defined on the CEFR scale, respectively. Given the students' similar level of L2 academic writing experience based on this information, the purpose to select the students' second set of term papers written in their fourth semester was to capture the developmental process by which they learn to employ rhetorical structure and linguistic features through the three-semester training of academic writing in a plausible way. After the three corpora were cleaned, with chapter titles, section headers, footers, equations, and graphics completely deleted, they were saved as electronic files. Corpus analysis software, *WordSmith Tools 7.0* (Scott, 2017), was used to provide two measures of lexical variety within a text, type/token ratio (TTR) and standardized type/token ratio (STTR). TTR refers to the ratio between the types and the tokens in any given text whereas STTR refers to the average type/token ratio based on analysis of consecutive 1,000-word chunks of text and is thus less sensitive to variations in text length. The description of the three corpora is shown in Table 1.

**TABLE 1**  
**The Composition of L1 Expert Corpus and Two L2 Student Corpora**

Corpus	No. Texts	No. Tokens	Average Words per Text	No. Types	TTR (%)	STTR (%)
English L1	200	124,511	623	9,791	7.86	41.95
English L2-2	49	31,572	644	3,626	11.48	37.42
English L2-1	46	25,629	557	3,456	13.48	40.17

### 3.2. Data Analysis

Since the purpose of this study was to discover the connection between the structural use of the identified lexical bundles and the communicative functions that they were conveying or helping convey in context, the methodology used in this study was bottom-up. Three sets of lexical bundles separately identified in a L1 corpus and two L2 corpora were analyzed in their context in order to examine the communicative function of discourse in proximity to each bundle and establish the rhetorical move and step. Due to many moves and steps which did not use lexical bundles in the three corpora, tagging for moves and steps was not necessary and the only stretches of discourse related to or in close proximity to the lexical bundles identified were examined for move analysis.

First, lexical bundles of four+ words were identified in that three-word bundles are frequently part of four-word bundles (Cortes, 2004). Frequency and dispersion were set as two important criteria in identifying lexical bundles. Even though previous bundle research

typically set normalized cut-off frequency and dispersion threshold at from 20 to 40 occurrences per million words (Biber et al., 2004; Chen & Baker, 2010) and 5 texts (Biber et al., 1999, 2004; Cortes, 2004, 2008, 2013; Pan, Reppen, & Biber, 2016), the corpora used in this study were not large enough to set this strict standard. Thus, the normalized cut-off frequency was set at 3 and 1 by calculating the ratio and the dispersion threshold was set at 3 and 2 considering corpus size for the L1 corpus and the two L2 corpora. Using the clusters/n-gram function of *AntConc* (Anthony, 2014), four+ word bundles in the three corpora were searched. Several context-dependent (e.g., *in the United States*) or topic-specific (e.g., *for second language learning*) bundles were excluded from the analysis (Chen & Baker, 2010; Huang, 2015). Following Chen and Baker (2010), the bundle overlaps including complete overlaps (e.g., *it has been suggested that including has been suggested that*) and complete subsumptions (e.g., *as a result of* and *a result of the* as a subset of *as a result of the*) were also manually checked in order to avoid inflating quantitative results. Subsequently, the extracted bundles were structurally categorized following the taxonomy in the Longman Grammar of Spoken and Written English (LSWE) (Biber et al., 1999) as shown in Table 2.

**TABLE 2**  
**Propositional Distribution of Lexical Bundles across the Structural Categories**  
**in the Three Corpora**

Patterns	Structural Categories
NP-based	1 noun phrase with of-phrase fragment (e.g., <i>the size of the</i> )
	2 noun phrase with other post-modifier fragment (e.g., <i>the difference between the</i> )
PP-based	3 prepositional phrase with embedded of-phrase (e.g., <i>in the case of</i> )
	4 other prepositional phrase fragment (e.g., <i>at the same of</i> )
VP-based	5 copula be + noun phrase/adjective phrase (e.g., <i>is due to the</i> )
	6 (verb phrase+) with active verb (e.g., <i>we can get the</i> )
	7 anticipatory it + verb phrase/adjective phrase (e.g., <i>it is easy to</i> )
	8 passive verb + prepositional phrase fragment (e.g., <i>can be seen in</i> )
	9 (verb phrase) + that clause fragment (e.g., <i>we assume that the</i> )
Others	10 (e.g., <i>as well as the</i> )

Next, move analysis was conducted to identify the moves/steps in which the lexical bundles were identified. Using corpus-based approach in which existing framework was used as a tool to validate linguistic categories (Tognini-Benelli, 2001), the L1 corpus was analyzed with a revised framework based on Swales's (1990, 2004) CARS framework with reference to several previous studies (Loi, 2010; Shehzad, 2008; Sheldon, 2011). However, the two L2 corpora were analyzed without any existing framework using a corpus-driven approach (McEnery, Xio, & Tono, 2006) based on the assumption that the rhetoric of the two L2 corpora is different from that of the L1 corpus. In this corpus-driven analysis, two

coders distinguished segments based on their global and local rhetorical purposes reading the texts carefully, categorized them into functional discourse units of analysis and then reduced the number of categories considering the issue of duplication. The functional units which are either in relative proximity to each other or often occur in similar locations were grouped in order to assign the specific steps that fulfill local rhetorical purposes and realize a broader move that achieves a global rhetorical purpose. The cutoffs for being a move/step and move stability or conventionality were set as 90/10 and 50 percent of the total RAs referring to prior research (Kanoksilapatham, 2007; Upton & Cohen, 2009). A final framework of moves/steps was developed after pilot-coding to test and fine-tune definitions of the tentative move/step schema.

In this study, only the context in which the lexical bundles were identified was analyzed in order to establish the rhetorical moves/steps and communicative functions. A second coder classified 50% of the randomly selected bundles identified and the percentage of agreement between two coders for communicative purpose identification was calculated. The inter-coder agreement was calculated with Cohen's kappa, and repeated after several negotiations, with a mean agreement of .88 for all bundles selected, which indicated high level of reliability. The steps of the three corpora that achieve 10 corresponding rhetorical functions in total are indicated with same figures (e.g., ●, □) as shown in Table 3.

**TABLE 3**  
**Steps that Have Corresponding Rhetorical Functions Between the Three Corpora**

English L1	English L2-2	English L2-1
○ Move 1-step 1 claiming centrality	Move 1-step 4 claiming centrality	Move 1-step 3 claiming centrality
● Move 1-step 2 making topic generalization(s) of increasing specificity	Move 1-step 1 presenting research topics and clarifying terms related to them Move 1-step 2 explaining research topics Move 1-step 3 drawing necessity of research area from educational situation Move 1-step 7 narrowing down and explaining research topics	Move 1-step 1 presenting research topics and clarifying terms related to them Move 1-step 2 explaining research topics Move 1-step 4 drawing necessity of research area from educational situation Move 1-step 7 narrowing down and explaining research topics
■ Move 1-step 3 reviewing items of previous research	Move 1-step 5 reviewing and summarizing previous research	Move 1-step 5 reviewing and summarizing previous research
□ Move 2-step 1A counter-claiming Move 2-step 1B	Move 1-step 6 question-raising Move 2-step 3	Move 1-step 6 question-raising Move 2-step 1

indicating a gap	addressing necessity of research	addressing necessity of research
Move 2-step 1C question-raising	Move 2-step 1 pointing out lack of research related to topics	Move 2-step 2 pointing out lack of research related to topics
Move 2-step 1D continuing a tradition	Move 2-step 2 presenting limitations of previous research	Move 2-step 3 presenting limitations of previous research
◆ Move 3-step 1 announcing present research descriptively and/or purposively	Move 3-step 1 presenting purposes of the present study	Move 3-step 1 presenting purposes of the present study
◇ Move 3-step 2 specifying the focus of the research Move 3-step 5 presenting positive justification	Move 3-step 2 explaining features, range, subjects, or premises of research	Move 3-step 2 explaining features, range, subjects, or premises of research
△ Move 3-step 3 foreshadowing the methodology	Move 3-step 3 explaining methodology or procedure of research	Move 3-step 3 explaining methodology or procedure of research
▲ Move 3-step 4 introducing the RQs or research hypothesis	Move 3-step 4 explaining research questions	Move 3-step 4 explaining research questions
▷ Move 3-step 7 introducing the implications of the findings	Move 3-step 5 presenting implications of research	Move 3-step 5 presenting implications of research
▶ Move 3-step 8 outlining the structure of the study	Move 3-step 6 presenting structure of research	Move 3-step 6 presenting structure of research

As represented in Table 3, in the three corpora, each rhetorical function of claiming centrality, reviewing or summarizing previous research, presenting purposes of the current study, explaining methods of the study, explaining research questions, presenting implications of the study, and presenting structure of the study was fulfilled in a step, indicated with ○, ■, ◆, △, ▲, ▷, and ▶. The step which achieved the function of making topic generalizations of increasing specificity in the L1 corpus corresponded to the 4 steps in the two L2 corpora with some differences in order, indicated with ●. The function of identifying the more specific areas of research that require further investigation was fulfilled with the 4 steps in the three corpora, indicated with □. The function of explaining the specific features of the research, which was found in the 2 steps in the L1 corpus was fulfilled in a step in the two L2 corpora, indicated with ◇.

## 4. RESULTS AND DISCUSSION

### 4.1. Propositional Distribution of Main Structural Categories of Lexical Bundles in the Three Corpora

The structural forms of the bundles of four+ words in the introductions of published RAs and those of two groups of student term papers were identified. As shown in Table 4 and Figure 1, significant differences in the main structural categories of lexical bundles (NP-based, PP-based and VP-based) across the three corpora were found. The developmental process of the Korean graduate students at both levels in the structural use of lexical bundles was evidenced by significant differences between the three corpora as well as between the two student corpora. The results shown in Table 4 also indicated that the use of more NP-based (e.g., *as one of the*) and PP-based bundles (e.g., *with the development of*) tended to increase going from the low to the high level and on to the expert level, while the use of VP-based bundles (e.g., *do not have the*) tended to decrease along the same line. The result is in line with several prior research in which L1 experts preferred the use of phrasal bundles consisting of NP-based and PP-based bundles, while L1/L2 novices and/or L2 experts showed more use of VP or clausal bundles (e.g., Biber, Staples, & Gray, 2016). This structural difference in bundle use between the three corpora reflects an important developmental level of academic writing featured by a shift from clausal style to phrasal style focusing on high information (Pan et al., 2016).

**TABLE 4**  
**Distribution of Structural Categories of Lexical Bundles in the Three Corpora**

	Types			Token		
	English L2-1	English L2-2	English L1	English L2-1	English L2-2	English L1
NP	44(33.6%)	49(31.4%)	59(35.5%)	150(39.7%)	255(47.1%)	209(46.4%)
PP	30(22.9%)	30(19.2%)	41(24.7%)	80(21.2%)	80(14.8%)	120(26.7%)
NP+PP	74(56.5%)	79(50.6%)	100(60.2%)	230(60.9%)	335(61.9%)	329(73.1%)
VP	31(23.7%)	53(34.0%)	45(27.1%)	78(20.6%)	133(24.6%)	80(17.8%)
Etc.	26(19.8%)	24(15.4%)	21(12.7%)	70(18.5%)	73(13.5%)	41(9.1%)
Total	131(100%)	156(100%)	166(100%)	378(100%)	541(100%)	450(100%)

Note. NP=NP-based phrasal, PP=PP-based phrasal, NP+PP= NP-based phrasal+PP-based phrasal,

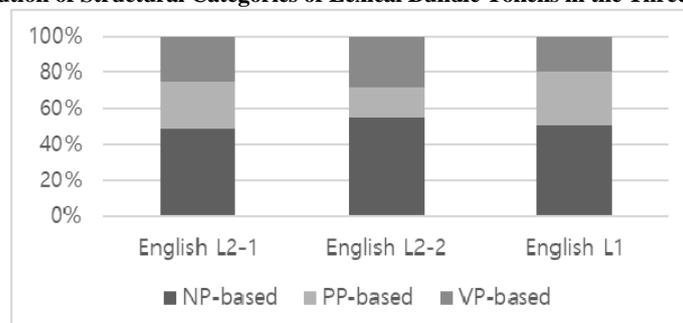
VP=VP-based phrasal, etc.=others

L2-1/L2-2/L1:  $\chi^2 = 39.192$ ,  $df = 6$ ,  $p = 0.000$

L2-1/L2-2:  $\chi^2 = 13.124$ ,  $df = 3$ ,  $p = 0.004$

L2-1/L1:  $\chi^2 = 19.182$ ,  $df = 3$ ,  $p = 0.000$

L2-2/L1:  $\chi^2 = 26.599$ ,  $df = 3$ ,  $p = 0.000$

**FIGURE 1****Distribution of Structural Categories of Lexical Bundle Tokens in the Three Corpora**

One noticeable finding was that the Korean graduate students at both levels showed increased use of both phrasal and clausal bundles, which are conflicting with the prior research. This finding can be interpreted due to the fact that the Korean graduate students are still in the process of increasing their bundle use, after which the developmental level of academic writing from clausal to phrasal takes place. An abrupt increase in the use of VP-based bundles in the high level students' corpus compared to the low level students (e.g., *should not be the*) also suggests the students' difficulty in the progression from clausal to phrasal style of academic writing or some specific clausal features used repetitively in their writing. According to this evidence gathered in the present study, the Korean graduate students at both levels are arguably in the stage that starts to show signs of transition, whereby they begin to grasp the distinction between formal and informal writing. Their level can be thus described as CEFR-B2 (the intermediate level) between CEFR-B1 (the lowest level) and CEFR-C1 (the highest level), which are clearly characterized by highly conversational and formal, respectively (Chen & Baker, 2016).

#### 4.2. Structural Distribution of Lexical Bundles in each Move in the Three Corpora

In order to systematically analyze the structural distribution of lexical bundles in each move in the three corpora from a L2 development perspective, all the bundles identified were classified according to the rhetorical moves of RA introductions. First, the distribution of the identified bundle tokens in different moves was presented in Tables 5–7. The findings showed a shift from clausal style using more VP-based bundles to phrasal style using more NP-based and PP-based bundles as the student level advanced from the low to the high level and on to the expert level, which are consistent with those of prior research (Chen & Baker, 2016; Pan et al., 2016; Qin, 2014; Ruan, 2017).

**TABLE 5**  
**Structural Distribution of Bundle Tokens in Move 1 in the Three Corpora**

	English L2-1	English L2-2	English L1
NP-based	24(20.00%)	31(23.66%)	116(26.85%)
PP-based	31(25.83%)	35(26.72%)	214(49.54%)
VP-based	65(54.17%)	65(49.62%)	102(23.61%)
Total	120(100%)	131(100%)	432(100%)

**TABLE 6**  
**Structural Distribution of Bundle Tokens in Move 2 in the Three Corpora**

	English L2-1	English L2-2	English L1
NP-based	3(8.33%)	7(23.33%)	39(22.16%)
PP-based	9(25.00%)	7(23.33%)	75(42.61%)
VP-based	24(66.67%)	16(53.33%)	62(35.22%)
Total	36(100%)	30(100%)	176(100%)

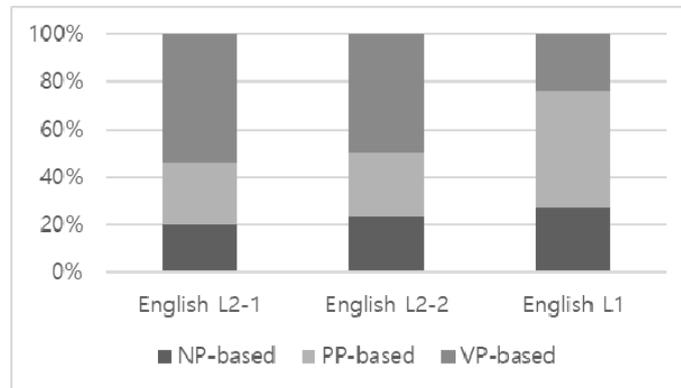
**TABLE 7**  
**Structural Distribution of Bundle Tokens in Move 3 in the Three Corpora**

	English L2-1	English L2-2	English L1
NP-based	11(50.00%)	23(51.11%)	42(28.77%)
PP-based	4(18.18%)	8(17.78%)	77(52.74%)
VP-based	7(31.82%)	14(31.11%)	27(18.49%)
Total	22(100%)	45(100%)	146(100%)

In addition, this shift in the structural use of lexical bundles reflected the Korean graduate students' developmental order of lexical bundles as well as their approximation to the norm of academic prose. That is, as shown in Figures 2–4, VP-based and NP-based bundles appeared to be acquired by the students prior to PP-based bundles, as evidenced by the increasing use of PP-based bundles at a much slower pace (e.g., *despite the fact that*) compared to the increasing use of NP-based bundles (e.g., *a wide range of*) and decreasing use of VP-based bundles (e.g., *are related to the*) as the student level advanced. This tendency was also confirmed in previous studies which investigated the use of lexical bundles in L2 students' academic writing across different proficiency levels or levels of studies (Chen & Baker, 2016; Ruan, 2017). In Chen and Baker's (2016) research, a similar pattern of bundle use across the three CEFR levels (B1, B2, and C1) of Chinese student writing was found. VP-based bundles were proportionally the biggest group in B1 level (the lowest level); NP-based bundles had a higher proportion in B2 than the other levels; and PP-based bundles had the widest range in C1 (the highest level). In the same vein, Ruan (2017) also evidenced later acquisition of PP-based bundles compared to the other bundles revealing much larger proportion of NP-based bundles in the lowest level student writing than in the other writings, only small variations of VP-based bundles across different levels of studies, and much higher proportions of PP-based bundles in the

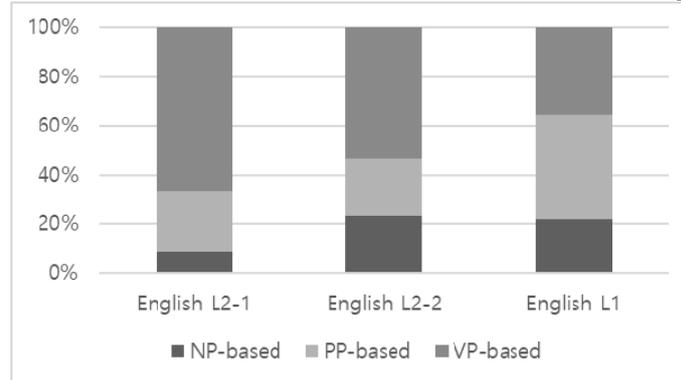
**FIGURE 2**

**Structural Distribution of Bundle Tokens in Move 1 in the Three Corpora**



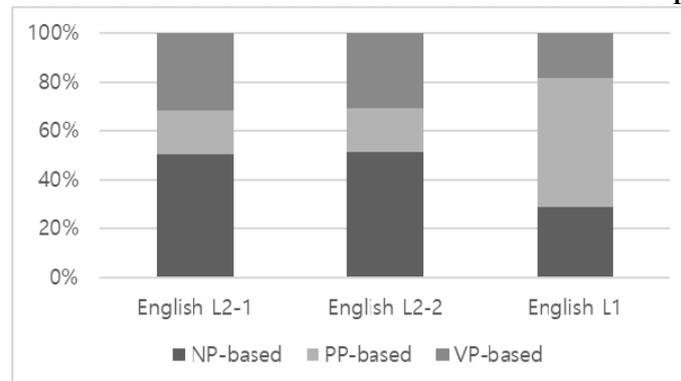
**FIGURE 3**

**Structural Distribution of Bundle Tokens in Move 2 in the Three Corpora**



**FIGURE 4**

**Structural Distribution of Bundle Tokens in Move 3 in the Three Corpora**



higher level student writing than the lower level student writing. The most frequent use of content-focused compound NP-based bundles was also evidenced as an indicator of the improvement on international students' academic literacy adaptation in Park's (2016) mixed-methods research. The reason for using this large portion of NP-based was assumed to be the particular academic genre type, critical review paper.

Next, all the quantitative analyses using chi-square tests were performed using the open source programming language, R, from the web, <https://www.r-project.org/>. The results of chi-square tests in Tables 8–10 indicated that no statistically significant differences were found between the low level students and the high level students in terms of the structural use of lexical bundles in Move 1, 2, and 3.

**TABLE 8**

**Standardized Residuals in a Chi-square Contingency Table for Structural Distribution of Bundle Tokens in Move 1 in the English L2-2 Corpus and the English L2-1 Corpus**

	Structural Category	NP-based	PP-based	VP-based
English L2-2	observed	31(23.66%)	35(26.72%)	65(49.62%)
	expected	28.70518	33.9243	67.8486
	R	0.43	0.09	-0.35
English L2-1	observed	24(20.00%)	31(25.83%)	65(54.17%)
	expected	26.29482	31.0757	62.15139
	R	-0.45	-0.1	0.36

Note.  $X^2 = 0.65$ ,  $df = 2$ ,  $p = 0.7225$ , cramer's  $V = 0.0509$

**TABLE 9**

**Standardized Residuals in a Chi-square Contingency Table for Structural Distribution of Bundle Tokens in Move 2 in the English L2-2 Corpus and the English L2-1 Corpus**

	Structural Category	NP-based	PP-based	VP-based
English L2-2	observed	7(23.33%)	7(23.33%)	16(53.33%)
	expected	4.54545	7.27273	18.18182
	R	1.15	-0.1	-0.51
English L2-1	observed	3(8.33%)	9(25.00%)	24(20.00%)
	expected	5.45454	8.72727	21.81818
	R	-1.05	0.09	0.47

Note.  $X^2 = 2.93$ ,  $df = 2$ ,  $p = 0.2311$ , cramer's  $V = 0.2107$

**TABLE 10**  
**Standardized Residuals in a Chi-square Contingency Table for Structural Distribution of Bundle Tokens in Move 3 In the English L2-2 Corpus and the English L2-1 Corpus**

	Structural Category	NP-based	PP-based	VP-based
English L2-2	observed	23(51.11%)	8(17.78%)	14(31.11%)
	expected	22.83582	8.0597	14.10448
	R	0.03	-0.02	-0.03
English L2-1	observed	11(50.00%)	4(18.18%)	7(31.82%)
	expected	11.16418	3.9403	6.89552
	R	-0.05	0.03	0.04

Note.  $X^2 = 0.01$ ,  $df = 2$ ,  $p = 0.995$ , cramer's  $V = 0.0122$

This suggests that the Korean graduate students in this study did not show noticeable evidence of L2 development in the structural use of lexical bundles during the transition from earlier to later academic level of graduate coursework. This finding is inconsistent with that of Qin's (2014) cross-sectional bundle study, in which noticeable difference in structural bundle use was found between the non-native graduate low-level students and high-level students. This can be presumably interpreted that simple exposure to academic writing through three-semester training did not necessarily transfer into the language used in the Korean graduate students' written output. Thus, more explicit practice in how to express information more economically by using more concise and compact structure with NP-based and PP-based bundles should be included in EAP writing courses.

Statistically significant differences were found between the low level students and the English experts in terms of the structural use of lexical bundles in Move 1, 2 and 3, as indicated in the results of chi-square tests in Tables 11–13. The calculated standardized residuals and absolute value greater than 1.96 in Table 11 and Table 12 showed that the cells in the contingency table indicating VP-bundles found in English L2-1 corpus contributed significantly to the observed differences. That is, the low-level students are challenged with approximating target constructions in academic writing, using much more VP-based bundles compared to the English experts, particularly in Move 1 and 2 as shown in text sample 1 and 2.

Text sample 1: L2-1 corpus

After that, what kinds of researches have done for using corpus for language instruction *is going to be* presents.

Text sample 2: L2-1 corpus

Despite those efforts, still, *it is unsure for* us to establish the best way to require it.

**TABLE 11****Standardized Residuals in a Chi-square Contingency Table for Structural Distribution of Bundle Tokens in Move 1 In the English L1 Corpus and the English L2-1 Corpus**

	Structural Category	NP-based	PP-based	VP-based
English L1	observed	116(28.85%)	214(49.54%)	102(23.61%)
	expected	109.56522	191.73913	130.69565
	R	0.61	1.61	-2.51
English L2-1	observed	24(20.00%)	31(25.83%)	65(54.17%)
	expected	30.43478	53.26087	36.30435
	R	-1.17	-3.05	4.76

Note.  $X^2 = 42.61$ ,  $df = 2$ ,  $p < 0.001$ , cramer's  $V = 0.2778$

**TABLE 12****Standardized Residuals in a Chi-square Contingency Table for Structural Distribution of Bundle Tokens in Move 2 In the English L1 Corpus and the English L2-1 Corpus**

	Structural Category	NP-based	PP-based	VP-based
English L1	observed	39(22.16%)	75(42.61%)	62(35.22%)
	expected	34.86792	69.73585	71.39623
	R	0.7	0.63	-1.11
English L2-1	observed	3(8.33%)	9(25.00%)	24(20.00%)
	expected	7.13207	14.26415	14.60377
	R	-1.55	-1.39	2.46

Note.  $X^2 = 12.51$ ,  $df = 2$ ,  $p = 0.0019$ , cramer's  $V = 0.2429$

**TABLE 13****Standardized Residuals in a Chi-square Contingency Table for Structural Distribution of Bundle Tokens in Move 3 In the English L1 Corpus and the English L2-1 Corpus**

	Structural Category	NP-based	PP-based	VP-based
English L1	observed	42(28.77%)	77(52.74%)	27(18.49%)
	expected	46.05952	70.39286	29.54762
	R	-0.6	0.79	-0.47
English L2-1	observed	11(50.00%)	4(18.18%)	7(31.82%)
	expected	6.94048	10.60714	4.45238
	R	1.54	-2.03	1.21

Note.  $X^2 = 9.15$ ,  $df = 2$ ,  $p = 0.0103$ , cramer's  $V = 0.2334$

The results of chi-square tests in Tables 14–15 revealed that statistically significant differences were found between the high level students and the English experts in terms of the structural use of lexical bundles in Move 1 and Move 3. The findings suggest that the even high-level students are not still able to employ appropriate bundle use in terms of structure in the same way that the English experts do. The calculated standardized residuals

and absolute value greater than 1.96 in Table 14 showed that the cells in the contingency table indicating VP-bundles found in English L2-2 corpus contributed significantly to the observed differences. That is, the high-level students still have difficulty in following the norm of academic writing, using much more VP-based bundles compared to the English experts, particularly in Move 1 as represented in text sample 3.

Text sample 3: L2-2 corpus

Some factors that relate to aural input and lack of background knowledge *have been found to* pose major obstacles to listening comprehension development.

**TABLE 14**

**Standardized Residuals in a Chi-square Contingency Table for Structural Distribution of Bundle Tokens in Move 1 In the English L1 Corpus and the English L2-2 Corpus**

	Structural Category	NP-based	PP-based	VP-based
English L1	observed	116(28.85%)	214(49.54%)	102(23.61%)
	expected	112.79574	191.06217	128.14209
	R	0.3	1.66	-2.31
English L2-2	observed	31(23.66%)	35(26.72%)	65(49.62%)
	expected	34.20426	57.93783	38.85790
	R	-0.55	-3.01	4.19

Note.  $X^2=35.15$ ,  $df=2$ ,  $p<0.001$ , cramer's  $V=0.2499$

One interesting finding gathered in the present study was found in the calculated standardized residuals and absolute value greater than 1.96 in Table 15, which indicated that NP-bundles found in English L2-2 corpus contributed significantly to the observed differences. The much more use of NP-based bundles in Move 3 by the high-level students compared to the English experts in this study is conflicting with the findings of previous literature (Chen & Baker, 2010, 2016). In this study, the NP-based bundles used in the Move 3-step 1 and Move 3-step 2 acted as triggers to achieve the specific communicative purpose as represented in text sample 4 and 5.

Text sample 4: L2-2 corpus

Stemmed from the issues discussed in the previous section, *the purpose of the present study is to* explore the language NNESTs use in their classrooms.

Text sample 5: L2-2 corpus

The automaticity is *one of the key* concepts of information processing theories, which explained the language acquisition as a process of linguistic information was learned, processed and stored in long term memory.

This result can be interpreted by the fact that the high level students tried to make frequent use of NP-based bundles directly extracting common noun phrases from published work in order to convey the particular rhetorical function of announcing the present study and pack their information in the most economical manner (Biber et al., 1999; Cortes, 2013). This language re-use phenomenon or language pervasiveness has been regarded as common practice in academic discourse (Flowerdew & Li, 2007).

**TABLE 15****Standardized Residuals in a Chi-square Contingency Table for Structural Distribution of Bundle Tokens in Move 3 In the English L1 Corpus and the English L2-2 Corpus**

	Structural Category	NP-based	PP-based	VP-based
English L1	observed	42(28.77%)	77(52.74%)	27(18.49%)
	expected	49.68586	64.97382	31.34031
	R	-1.09	1.49	-0.78
English L2-2	observed	23(51.11%)	8(17.78%)	14(31.11%)
	expected	15.31414	20.02618	9.65968
	R	1.96	-2.69	1.4

Note.  $X^2 = 17.05$ ,  $df = 2$ ,  $p = 0.0002$ , cramer's  $V = 0.2988$

However, the result of a chi-square tests in Table 16 indicated that no statistically significant difference was found between the high-level students and the English experts in terms of the structural use of lexical bundles in Move 2. This suggests that the high-level students are able to employ the lexical bundles in terms of structure in the appropriate way the English experts do in Move 2. That is, the Korean graduate students tend to acquire the structural use of lexical bundles later in Move 1 and 3 compared to in Move 2.

**TABLE 16****Standardized Residuals in a Chi-square Contingency Table for Structural Distribution of Bundle Tokens in Move 2 In the English L1 Corpus and the English L2-2 Corpus**

	Structural Category	NP-based	PP-based	VP-based
English L1	observed	39(22.16%)	75(42.61%)	62(35.22%)
	expected	39.30097	70.05825	66.64078
	R	-0.05	0.59	-0.57
English L2-2	observed	7(23.33%)	7(23.33%)	16(53.33%)
	expected	6.69903	11.94175	11.35922
	R	0.12	-1.43	1.38

Note.  $X^2 = 4.63$ ,  $df = 2$ ,  $p = 0.0988$ , cramer's  $V = 0.1499$

## 5. CONCLUSION

The study aimed to investigate the features of the process by which Korean graduate students at two different academic levels have acquired the structural use of lexical bundles in the rhetorical moves compared to English experts, particularly in RAs as a specific genre.

Major findings of the study are as follows. First, as for propositional distribution of main structural categories of lexical bundles (NP-based, PP-based and VP-based), significant differences between the three corpora as well as between the two student corpora suggest that the Korean graduate students are in the developmental process of academic writing featured by a shift from clausal style using more VP-based bundles to phrasal style using more NP-based and PP-based bundles as their academic level advances (Biber, Staples, & Gray, 2016; Pan et al., 2016; Qin, 2014). Second, the structural distribution of lexical bundles in each move in the three corpora also confirms the students' developmental order of lexical bundles that VP-based and NP-based bundles tend to be acquired prior to PP-based bundles (Chen & Baker, 2016; Ruan, 2017) in addition to the approximation to the norm of academic prose. Next, the results of chi-square tests of the three corpora suggest that the Korean graduate students in this study did not show noticeable development in the structural use of lexical bundles even during the transition from earlier to later academic level of graduate coursework. The last noticeable finding was that even the high-level students still had difficulty in following the norm of academic writing, using much more VP-based bundles in Move 1 and much more NP-based in Move 3 compared to the English experts.

The current study lends empirical support to the claim that the Korean graduate students as EAP learners are in the developmental process of acquiring the bundle use in terms of structure. By pointing out that the high-level students were not able to use bundles in the same way that the English experts did in terms of structure, this study also suggests explicit teaching is required to accelerate these EAP students' acquisition of target lexical bundles.

In closing, the findings of this study have pedagogical implications for teaching academic writing in EAP contexts. The developmental patterns of specific structural category of bundles in different moves found in L2 student writing can inform the design of an academic writing course for EAP students who are required to construct disciplinary knowledge for communicating in the academic discourse community. Supporting the view that the instruction of lexical bundles should be integrated into the EAP curriculum (Chen & Baker, 2010; Hyland, 2008a, 2008b), it is thus recommended that EAP students even at earlier level of graduate coursework be exposed to structural frames (e.g., *the NOUN + of the/a*) and examples of bundles (e.g., *a better understanding of the, our knowledge of the, the results of the*) for each structural category (NP-based, PP-based and VP-based) typically used in each move considering the students' acquisition order of structural bundle

use in the specific move. Further follow-up instructional studies in which the target lexical bundles are taught explicitly and more longitudinal analysis with large-scale corpora reliably annotated at the step level than at the move level (Cotos et al., 2017; Flowerdew, 2002) can provide more interesting and robust findings with regard to L2 development.

## REFERENCES

- Anthony, L. (2014). *AntConc 3.4.3* [Computer Software]. Tokyo, Japan: Waseda University. Retrieved on March 03, 2019, from the World Wide Web: <http://www.aurenceanthony.net/>
- Ädel, A., & Erman, B. (2012). Recurrent word combinations in academic writing by native and non-native speakers of English: A lexical bundles approach. *English for Specific Purposes, 31*(2), 81-92.
- Biber, D., & Barbieri, F. (2007). Lexical bundles in university spoken and written registers. *English for Specific Purposes, 26*(3), 263-286.
- Biber, D., Connor, U., & Upton, T. (2007). *Discourse on the move: Using corpus analysis to describe discourse structure*. Amsterdam: John Benjamins.
- Biber, D., Conrad, S., & Cortes, V. (2004). If you look at.: Lexical bundles in university teaching and textbooks. *Applied Linguistics, 25*(3), 371-405.
- Biber, D., Conrad, S., & Leech, G. (2002). *Longman student grammar of spoken and written English*. Harlow: Pearson.
- Biber, D., Johansson, S., Leech, G., Conrad, S., & Finegan, E. (1999). *Longman grammar of spoken and written English*. Harlow: Pearson.
- Biber, D., Staples, S., & Gray, B. (2016). Predicting patterns of grammatical complexity across language exam task types and proficiency levels. *Applied Linguistics, 37*(5), 639-668.
- Chen, Y. H., & Baker, P. (2010). Lexical bundles in L1 and L2 academic writing. *Language Learning & Technology, 14*(2), 30-49.
- Chen, Y. H., & Baker, P. (2016). Investigating criterial discourse features across second language development: Lexical bundles in rated learner essays, CEFR B1, B2, and C1. *Applied Linguistics, 37*(6), 849-880.
- Connor, U., Davis, J., & De Rycker, T. (1995). Correctness and clarity in applying for overseas jobs: a cross-cultural analysis of US and Flemish applications. *Text, 15*(4), 457-475.
- Connor, U., & Upton, T. (2004). *Discourse in the professions: Perspectives from corpus linguistics*. Amsterdam: John Benjamins.
- Conrad, S., & Biber, D. (2004). The frequency and use of lexical bundles in conversation

- and academic prose. *Lexicographica: International Annual for Lexicography*, 20, 56-71.
- Cortes, V. (2004). Lexical bundles in published and student disciplinary writing: Examples from history and biology. *English for Specific Purposes*, 23(4), 397-423.
- Cortes, V. (2008). A comparative analysis of lexical bundles in academic history writing in English and Spanish. *Corpora*, 3(1), 43-58.
- Cortes, V. (2013). The purpose of this study is to: Connecting lexical bundles and moves in RA introductions. *Journal of English for Academic Purposes*, 12(1), 33-43.
- Cotos, E. (2014). *Genre-based automated writing evaluation for L2 research writing: From design to evaluation and enhancement*. New York: Palgrave. Macmillan.
- Cotos, E., Haufman, S., & Link, S. (2017). A move/step model for methods sections: Demonstrating rigour and credibility. *English for Specific Purposes*, 46, 90-106.
- Council of Europe. (2001). *Common European Framework of Reference for Languages: Learning, teaching, assessment*. Cambridge University Press.
- Coxhead, A., & Byrd, P. (2007). Preparing writing teachers to teach the vocabulary and grammar of academic prose. *Journal of Second Language Writing*, 16(3), 129-147.
- De Cock, S. (2000). Repetitive phrasal chunkiness and advanced EFL speech and writing. In C. Mair & M. Hundt (Eds.), *Corpus linguistics and linguistic theory* (pp. 51-68). Amsterdam: Rodopi.
- Dudley-Evans, T. (1995). Genre: how far can we, should we go? *World Englishes*, 16(3), 353-358.
- Durrant, P. (2017). Lexical bundles and disciplinary variation in university students' writing: Mapping the territories. *Applied Linguistics*, 38(2), 165-193.
- Durrant, P., & Mathews-Aydnli, J. (2011). A function-first approach to identifying formulaic language in academic writing. *English for Specific Purposes*, 30(1), 58-72.
- Ellis, N. (2002). Frequency effects in language processing: A review with implications for theories of implicit and explicit language acquisition. *Studies in Second Language Acquisition*, 24(2), 143-188.
- Ellis, N. (2006). Cognitive perspectives on SLA. The associative cognitive CREED. *Association Internationale de Linguistique Appliquée Review*, 19(1), 100-121.
- Flowerdew, L. (2002). Genre in the classroom: A linguistic perspective. In A. M. Johns (Ed.), *Genre in the classroom: Multiple perspectives* (pp. 91-104). Mahwah, NJ: Lawrence Erlbaum.
- Flowerdew, L. (2009). Applying corpus linguistics to pedagogy: A critical evaluation. *International Journal of Corpus Linguistics*, 14(3), 393-417.
- Flowerdew, J., & Li, Y. (2007). Language re-use among Chinese apprentice scientists writing for publication. *Applied Linguistics*, 28(3), 440-465.

- Grabowski, L. (2015). Keywords and lexical bundles within English pharmaceutical discourse: A corpus-driven description. *English for Specific Purposes*, 38, 23-33.
- Guinda C. S. (2015). Genres on the move: Currency and erosion of the genre moves construct. *Journal of English for Academic Purposes*, 19, 73-87.
- Henry, A., & Rosenberry, R. L. (2001). A narrow-angled corpus analysis of moves and strategies of the genre: 'Letter of Application'. *English for Specific Purposes*, 20(2), 153-167.
- Huang, K. (2015). More does not mean better: Frequency and accuracy of lexical bundles in Chinese EFL learners' essay writing. *System*, 53, 12-23.
- Hyland, K. (2007). Applying a gloss: Exemplifying and reformulating in academic discourse. *Applied Linguistics*, 28(2), 266-285.
- Hyland, K. (2008a). As can be seen: lexical bundles and disciplinary variation. *English for Specific Purposes*, 27(1), 4-21.
- Hyland, K. (2008b). Academic clusters: text patterning in published and postgraduate writing. *International Journal of Applied Linguistics*, 18(1), 41-62.
- Hyland, K. (2013). Faculty feedback: perceptions and practices in L2 disciplinary writing. *Journal of Second Language Writing*, 22, 240-53.
- Hyland, K. (2015). *Academic publishing: Issues and challenges in the construction of knowledge*. Oxford, UK: Oxford University Press.
- Hyland, K. (2016). Methods and methodologies in second language writing research. *System*, 59, 116-125.
- Kanoksilapatham, B. (2003). *A corpus-based investigation of scientific research articles: Linking move analysis with multidimensional analysis*. Unpublished doctoral dissertation, Georgetown University, Washington, DC.
- Kanoksilapatham, B. (2005). Rhetorical structure of biochemistry research articles. *English for Specific Purposes*, 24(3), 269-292.
- Kanoksilapatham, B. (2007). Rhetorical moves in biochemistry RAs. In D. Biber, U. Connor, & T. A. Upton (Eds.), *Discourse on the move: Using corpus analysis to describe discourse structure* (pp. 73-119). Amsterdam: John Benjamins.
- Le, T. N. P., & Harrington, M. (2015). Phraseology used to comment on results in the discussion section of applied linguistics quantitative research articles. *English for Specific Purposes*, 39, 45-61.
- Lim, J. M. H. (2006). Method sections of management research articles: A pedagogically motivated qualitative study. *English for Specific Purposes*, 25(3), 282-309.
- Loi, C. K. (2010). Research article introductions in Chinese and English: A comparative genre-based study. *Journal of English for Academic Purpose*, 9(4), 267-279.
- Loi, C. K., & Lim, J. M. H. (2013). Metadiscourse in English and Chinese RA introductions. *Discourse Studies*, 15(2), 129-146.

- McEnergy, T, Xiao, R, & Tono, Y. (2006). *Corpus-based language studies: An advanced resource book*. New York: Routledge.
- Mizumoto, A. (2016). *Developing a writing support tool "AWSuM": Challenges and pedagogical applications*. Paper presented at the American association for corpus linguistics. Ames, Iowa: Iowa State University.
- Mizumoto, A., Hamatani, S., & Imao, Y. (2017). Applying the bundle-move connection approach to the development of an online writing support tool for research articles. *Language Learning*, 67(4), 885-921.
- Moreno, A. I. (2008). The importance of comparable corpora in cross-cultural studies. In U. Connor, E. Nagelhout, & W. V. Rozycki (Eds.), *Contrastive rhetoric: Reaching to intercultural rhetoric* (pp. 147-168). Amsterdam: John Benjamins.
- Moreno, A. I., & Swales, J. M. (2018). Strengthening move analysis methodology towards bridging the function-form gap. *English for Specific Purposes*, 50, 40-63.
- Myles, F., Hooper, R., & Mitchell, R. (1998). Rote or rule? Exploring the role of formulaic language in classroom foreign language learning. *Language Learning*, 48(3), 323-363.
- Omidian, T., Shahriari, H., & Siyanova-Chanturia, A. (2018). A cross-disciplinary investigation of multi-word expressions in the moves of research article abstracts. *Journal of English for academic purposes*, 36, 1-14.
- Paltridge, B. (2013). Genre and English for specific purposes. In B. Paltridge & S. Starfield (Eds.), *The handbook of English for specific purposes* (pp. 347-366). New York: Wiley-Blackwell.
- Pan, F., Reppen, R., & Biber, D. (2016). Comparing patterns of L1 versus L2 English academic professionals: Lexical bundles in telecommunications research journals. *Journal of English for Academic Purposes*, 21, 60-71.
- Park, E. (2016). Issues of international students' academic adaptation in the ESL writing class: A mixed-methods study. *Journal of International Students*, 6(4), 887-904.
- Qin, J. (2014). Use of formulaic bundles by non-native English graduate writers and published authors in applied linguistics. *System*, 42, 220-231.
- Ruan, Z. (2017). Lexical bundles in Chinese undergraduate academic writing at an English medium university. *RELC Journal*, 48(3), 327-340.
- Salazar, D. (2014). *Lexical bundles in native and non-native scientific writing: Applying a corpus-based study to language teaching*. Philadelphia, PA: John Benjamins.
- Scott, M. (2017). *WordSmith tools version 7.0*. Stroud, England: Lexical analysis software. Retrieved on March 01, 2019, from the World Wide Web: <http://www.lexically.net/wordsmith/version7/>
- Shehzad, W. (2008). Move two: Establishing a niche. *Ibérica*, 15, 25-50.
- Sheldon, E. (2011). Rhetorical differences in RA introductions written by English L1 and

- L2 and Castilian Spanish L1 writers. *Journal of English for Academic Purposes*, 10(4), 238-251.
- Simpson-Vlach, R., & Ellis, N. C. (2010). An academic formulas list: New methods in phraseology research. *Applied Linguistics*, 31(4), 487-512.
- Staples, S., Egbert, J., Biber, D., & McClair, A. (2013). Formulaic sequences and EAP writing development: Lexical bundles in the TOEFL iBT writing section. *Journal of English for Academic Purposes*, 12(3), 214-225.
- Swales, J. M. (1981). *Aspects of article introduction*. Birmingham, UK: University of Aston Language Studies Unit.
- Swales, J. M. (1990). *Genre analysis: English in academic and research settings*. Cambridge: Cambridge University Press.
- Swales, J. M. (2004). *Research genres: Exploration and applications*. Cambridge: Cambridge University Press.
- Tognini-Bonelli, E. (2001). *Corpus linguistics at work*. Amsterdam: John Benjamins.
- Upton, T., & Cohen, M. A. (2009). An approach to corpus-based discourse analysis: The move analysis as example. *Discourse Studies*, 11(5), 585-605.
- Wei, Y., & Lei, L. (2011) Lexical bundles in the academic writing of advanced Chinese EFL learners. *RELC Journal*. 42(2), 155-166.
- Wray, A. (2002). *Formulaic language and the lexicon*. Cambridge: Cambridge University Press.
- Yang, R., & Allison, D. (2003). Research articles in applied linguistics: Moving from results to conclusions. *English for Specific Purposes*, 22(4), 365-385.

Applicable levels: Tertiary

Ji-yoon Hong  
Lecturer  
Department of English Education  
College of Education, Hankuk University of Foreign Studies  
107 Imun-Ro, Dongdaemun-Gu  
Seoul 02450, Korea  
Email: my-sun6@hanmail.net

Received on July 27, 2019

Reviewed on September 1, 2019

Revised version received on September 5, 2019