

## **Cohesive Devices in English Writing Textbooks and Korean Learners' English Writings**

**Hee Youn Cho**

(Seoul National University)

**Jeong-Ah Shin\***

(Kwangwoon University)

**Cho, Hee Youn, & Shin, Jeong-Ah. (2014). Cohesive devices in English writing textbooks and Korean learners' English writings. *English Teaching*, 69(1), 41-59.**

The purpose of this study is to examine cohesive devices in English writing textbooks and Korean college students' writings through text and corpus analyses. An analysis of three writing textbooks used in one college English program in Seoul showed that the textbooks did not cover a full range of cohesive devices; instead, they focused on sentence transitions and conjunctions, which L2 writers often overuse or misuse. Other cohesive devices such as demonstratives and lexical cohesive devices such as synonyms, paraphrasing, or collocations, however, were rarely covered. To understand how Korean college students actually use cohesive devices in writing, this study also analyzed the frequencies of sentence transitions and demonstratives in learner and native speaker corpora. The results revealed L2 learners' tendency to overuse sentence transitions and demonstrative pronouns compared to native speakers. However, the results also showed that as proficiency increases, learners tend to use fewer sentence transitions.

**Key words:** cohesive devices, English academic writing, English academic writing textbooks, corpus analyses, text analyses

### **1. INTRODUCTION**

English has become the dominant language in the academic domain throughout the world (Ammon, 2001), and the preponderance of English can be easily witnessed in academic publications and conference presentations. For example, more than 90 percent of journal articles are written in English in scientific fields, and this proportion is even higher

---

\* Hee Youn Cho: First author; Jeong-Ah Shin: Corresponding author

for prestigious and most-cited journals (Hyland, 2006). The situation is not much different in Korea. English proficiency has become essential for students to be able to navigate their disciplines, exchange knowledge, and build future careers.

Although all spectra of English skills are important, helping and enabling students to produce papers and reports in English has emerged as one of the important goals of English programs in higher education. Producing technical texts that are linguistically adequate and that meet the discourse conventions of the particular field is an especially demanding task for non-native speakers of English. To gain an understanding of second language (L2) writing processes and to provide implications for L2 writing pedagogy, a line of research on writing in English as a Second Language (ESL) or English as a Foreign Language (EFL) has identified and compared significant differences among features of texts written by first language (L1) users and L2 users (Hinkel, 2011).

Previous research has documented numerous differences between L1 and L2 texts at the micro level, which can be summarized as limited vocabulary and grammar (Hinkel, 2004, 2011). At the macro level, global properties such as discourse organization, argumentation, and coherence/cohesion of L2 texts have shown to diverge from those of L1 texts (Choi, 1988; Hinkel, 2011). When writers produce longer stretches of text beyond the levels of clauses or sentences, the macro or discourse level knowledge and skills are essential as well as vocabulary and grammar knowledge. Although the research in L2 writing has significantly advanced the understanding of the characteristics of L2 discourse structuring, much still remains to be explored concerning how L2 discourse organization patterns develop as L2 writers gain more experience in the L2. Also, a gap exists between what accumulated research findings on L2 discourse development have suggested to be problematic and what areas L2 writing instruction has most focused on.

The present study aims to bridge these gaps, focusing on coherence/cohesion, one of the macro or discourse level features. In section 2, previous studies on coherence/cohesion will be reviewed. Section 3 presents an analysis of widely-used writing textbooks to examine how they approach coherence/cohesion. Section 4 presents corpus analyses of writing samples produced by Korean EFL university students to investigate how their coherence/cohesion strategies differ from those used by L1 writers and what role their English proficiency plays. Section 5 summarizes the general findings of the paper and discusses the implications for L2 writing instruction.

## 2. LITERATURE REVIEW

### 2.1. Coherence and Cohesion

Effective writing depends on more than well-formedness and clarity of individual sentences. In effective writing, each sentence easily connects to the previous and following sentences, and the text constitutes a unified whole. The textual properties that make text connected or coherent are referred to as cohesion (Halliday & Hasan, 1976). Various lexical, grammatical, and semantic means contribute to text cohesion, among which Halliday and Hasan (1976) discuss five major categories: reference, substitution, ellipsis, conjunction, and lexical cohesion.

The term reference indicates the relationship between two linguistic expressions. Pronominal reference (*he, they, him, etc.*), demonstratives (*this, these, that, those, here, there, etc.*), and comparative reference (*another, same, similar, such, so, etc.*) fall in this category. Substitution and ellipsis are ways of avoiding repetition of a lexical item. In substitution, one item is replaced with another, such as *one(s)* in nominal, *do (so)* or *does (so)* in verbal, and *so* in clausal substitution. In ellipsis, an item is elided, and it includes three subtypes: nominal, verbal, and clausal. Next, conjunction signals a relationship between parts of the discourse by linking them. Conjunctions can be further classified into four categories: additive (*and, furthermore, additionally, too, etc.*), adversative (*but, rather, yet, on the other hand, etc.*), causal (*therefore, so, as a result, for this reason, etc.*), and temporal (*then, next, first, in conclusion, etc.*). Lastly, lexical cohesion refers to bringing cohesive effect through the selection of vocabulary. There are five subcategories of lexical cohesion: repeating the same word or phrase, using a synonym, using a superordinate term, using a general word, and using lexical items that are likely to occur together, or collocation (Halliday & Hasan, 1976).

In addition to these five devices originally discussed in Halliday and Hasan (1976), researchers have identified other linguistic means that contribute to text cohesion. For example, Hinkel (2001) views nouns associated with categorization and division (e.g., *class, type, category, etc.*) as one of the important cohesive devices. Campbell (1995) points out that the continuity of verb tense and aspect is one of the key features of cohesive text. These cohesive devices help bind elements of a text together. In addition, Halliday and Hasan (1976) claimed that they also contributed to the coherence of a text, the property of a text making it a focused and united whole. While cohesion may contribute to the coherence of a text, other researchers claim that the two should not be equated (Carrell, 1982; Hinkel, 2001). Carrell (1982) points out that coherence is not achieved through mere surface linguistic features in the text. Rather, coherence is argued to be a content matter which is construed at the underlying propositional level.

The theories of cohesion and coherence have turned the attention of text analysis from sentence level features to encompass important inter-sentential and discourse-level flow and connections. They also have greatly affected ESL/EFL writing instruction; coherence and cohesion have become one of the key instructional objectives in L2 writing textbooks. According to Carrell's (1982) observation, however, some textbooks treat the concept of coherence on par with surface cohesion. Those textbooks reflect the unwarranted belief that teaching cohesion to students can make their writing more coherent. Similarly, Hinkel (2001) notes that L2 instruction on coherence and cohesion tends to focus on only a subset of linguistic means such as sentence transitions and coordinating conjunctions, as they can be more overtly taught than other subtle devices like using synonyms or paraphrasing a word or idea. It would be worth investigating how instruction practice has advanced in this area since the publication of Carrell's (1982) and Hinkel's (2001) research.

## 2.2. Coherence and Cohesion in L1 vs. L2 Texts

Much research has examined how L2 writers achieve coherence and cohesion in comparison to L1 writers. Research on cohesion has been especially fruitful as cohesive devices can be explicitly traced or observed on the surface. What has emerged out of this line of research is that L2 writers often underuse, overuse, and misuse certain types of cohesive devices (Bolton, Nelson, & Hung, 2003; Granger & Tyson, 1996; Hinkel, 2001; Kang, 2009).

Bolton et al. (2003) and Granger and Tyson (1996) focused on the use of sentence transitions such as *however*, *therefore*, and *for example*. Bolton et al. (2003) analyzed and compared the writing samples of university students in Hong Kong and in Great Britain. The frequency of sentence transitions in professional academic writing was used as the reference corpus to determine the extent to which the students overused or underused sentence transitions. The results showed that not only nonnative speakers but also native speakers overused a wide range of sentence transitions. There was no evidence of significant underuse of sentence transitions.

Granger and Tyson (1996) analyzed a large corpus of advanced EFL learner writing known as ICLE (International Corpus of Learner English) focusing on argumentative essays written by advanced learners of English who spoke French as their first language. A comparable control corpus containing essays written by native speakers was assembled and analyzed. The results showed that the learners tended to overuse sentence transitions of corroboration and giving examples, and adding points, while they tended to underuse transitions that signal contrast or develop the argument. In terms of the syntactic position, the learners significantly preferred the sentence initial position for connectors. The study discussed the L1 effects and insufficient knowledge of semantic, stylistic, and syntactic

restrictions on individual transitional expressions. Bolton et al. (2003) and Granger and Tyson (1996), however, are limited in scope, in that they only analyzed sentence transitions without reference to other kinds of cohesive devices.

Other studies focused on different sub-groups of cohesive devices, including contrastive conjunctions such as *but* and *yet* (e.g., Park, 2013), conjunctive adjuncts including conjunctions and conjunctive adverbials (e.g., Yoon & Yoo, 2011), and reference (e.g., Kang, 2009; Kim, 2012). Park (2013) examined Korean college EFL students' use of contrastive conjunctions in argumentative essays and found that the nonnative speakers overused and underused certain contrastive conjunctions compared to the native control group. They also tended to place contrastive conjunctions in sentence-initial position. The overall patterns were not very different in high proficiency and low proficiency groups.

Yoon and Yoo (2011) focused on grammatical errors associated with conjunctive adjuncts in Korean college students' essays. The learners were found to use sentence-initial coordinators even when they were not warranted, frequently produce sentence fragments with conjunctive adjuncts, and lack full knowledge of punctuation after conjunctive adjuncts. Kim (2012) investigated how Korean EFL learners' referential strategies differed from those of native speakers. The Korean EFL students, regardless of their proficiency level, used the pronoun *it* to refer to an extended set of text in the preceding discourse, whereas the pronoun produced by the native group most frequently referred to a nominal item. Also, the learners showed less mastery of using referential expressions in positions other than subject positions. Kang (2009), through the analysis of written narratives, showed that Korean EFL learners relied heavily on nominals than pronouns to establish textual cohesion.

Hinkel (2001) provided a more comprehensive analysis of the kinds and frequencies of cohesive devices used in academic essays written by 895 native and 697 nonnative writers. The nonnative group consisted of Japanese, Korean, Indonesian, and Arabic L1 speakers; they were all fairly proficient in English, with a mean TOEFL score of 587. Various kinds of cohesive devices were analyzed, including phrase-level coordinators (e.g., *also*, *and*), sentence transitions (e.g., *first*, *next*, *in sum*), logical and semantic conjunctions (e.g., *because of*, *in contrast*), demonstrative pronouns (e.g., *this*, *that*), enumerative nouns (e.g., *fact*, *factor*, *problem*), and resultative nouns (e.g., *effect*, *result*).

Overall, the nonnative writers used sentence transitions significantly more frequently than the native writers regardless of their L1. In the Korean learner group, the frequency of sentence transitions was more than three times higher than in the native speaker group. Many instances of sentence transitions that these non-native speakers used were not appropriate or effective in the given context. Hinkel attributed the results to overemphasis of sentence transition devices in L2 writing instruction. Another notable group difference was found in the use of demonstratives. With the exception of Indonesian speakers, the

learner groups used twice to three times as great demonstrative pronouns as the native speakers. In English, demonstrative pronouns function as effective cohesive markers only when identifiable referents can be found in proximity in the text. The learners, however, often used demonstrative pronouns to refer to broader contexts or ideas without specific textual referents. The study underscores the importance of expanding the repertoire of cohesive devices for learners who have a limited range of lexical and syntactic means of constructing a united flow. The study is quite comprehensive in that it includes learners from four different L1 backgrounds, but it does not show how these learners' use of cohesive devices develop as they gain more experience in the L2 or as their proficiency improves.

### 2.3. Corpus-based Analysis in Academic Writing

Corpus linguistics has provided a means for the empirical analysis of language, leading to a proliferation of empirical studies about aspects of grammar as well as the refinement of descriptions of lexis and enhanced coverage in dictionaries (McCarthy & O'Keeffe, 2010). Recently, corpus linguistics has been used in pursuit of broader research questions related to language teaching and learning (e.g., Aston, 1997; Oh, 2007, 2009). In the area of language teaching and learning (especially in teaching English as a second/foreign language), researchers have two different goals when conducting corpus-based studies: using native speaker corpora as authentic input for L2 learners, and creating L2 learner corpora and comparing L2 learner writing with native speaker norms.

Concerning the first goal, using native speakers' corpora and corpus analysis tools (e.g., concordancers) can be an effective method to learn and study languages. Native speaker corpora can provide L2 learners with a rich and extensive data source of examples of native speaker production, which can serve as authentic input for L2 learners. Also, since data from native speakers' corpora (most typically concordances) can be used in a hands-on manner (McCarthy & O'Keeffe, 2010), L2 learners in classrooms can experience data-driven learning (Johns, 1991), where learners can look at examples and the examples can help L2 learners generalize structures of sentences or usage of particular words. Thus, teachers can provide explicit instruction to L2 learners by showing them authentic data—that is, a corpus—and by letting them study native speakers' language patterns by themselves.

Second, corpus-based analyses aim to find differences between L2 learner writing and native speaker norms by comparing L2 learner corpora with native speaker corpora. Based upon the different patterns, common errors or mistakes that L2 learners make in comparison with native speakers can be highlighted. That is, studies of learner corpora aim to analyze learners' interlanguage in their writing and then demonstrate how their writing

differs from native speakers' writing. This can be closely related to the analysis of learner errors, so-called Error Analysis, which was prevalent in second language research during the 1960s and 1970s, but corpus-based analyses go beyond what Error Analysis could do. Error Analysis was criticized because it had no means of detecting L2 learners' avoidance behavior (i.e., avoiding difficult structures in their production; Ellis, 1994), but corpus-based analysis can overcome the limitations of Error Analysis by "comparing the frequency of words/structures in learner and native corpora" (Granger, 1998, p. 6), and thereby detect instances of avoidance. Therefore, using and analyzing native and learner corpora can shed light on language studies as a new method of pedagogy.

Using L2 learners' corpora, this study investigated L2 learner writing, focusing on the use of cohesive devices. Before reporting the results from corpus analyses, the study first presents a descriptive analysis of how L2 writing textbooks approach matters of cohesion. The specific research questions of the present study are:

1. How does current L2 pedagogy deal with text cohesion?
2. Does the use of cohesive devices differ by L2 proficiency level in writing?

### 3. TEXTBOOK ANALYSES

As discussed in the previous section, past L2 instruction or textbooks were criticized for equating coherence with using explicit cohesive expressions and for focusing on limited types of cohesive devices, such as sentence transitional expressions and conjunctions (Carrell, 1982; Hinkel, 2001). To determine whether the current L2 writing textbooks are subject to the same criticism, the study provides a brief description of how academic writing textbooks deal with coherence and cohesion. Three writing textbooks that were either used in the past or are currently being used in an English writing program for high-intermediate students (whose TEPS scores range from 701-800 or who have completed a lower level English course) at one university in Seoul were selected and analyzed based on Halliday and Hasan's (1976) classification of cohesive devices. The three textbooks analyzed in the study are *Engaging Writing 1* (Fitzpatrick, 2011), *Great Writing 3* (Folse, Solomon, & Clabeaux, 2010), and *Introduction to Academic Writing* (Oshiman & Hougue, 2007).

To report the findings, Fitzpatrick (2011) does not mention the terms coherence and cohesion but states that "the sentences in an effective paragraph must be clearly connected to one another so that readers can follow the ideas easily" (p. 106). It introduces three ways of making sentences connected, or achieving cohesion: using transitions, repeated words, and related words. In each chapter, sentence transitions and conjunctions for different

paragraph types such as cause-effect and contrast are introduced. However, lexical cohesion (using repeated words or related words) is illustrated in one simple exercise question. Some pages are devoted to exercises on varying word forms, but they all involve word- or sentence-level pattern drills without reference to how formally-related words contribute inter-sentential cohesion.

Folse et al. (2010) state that “a piece of writing is said to have coherence when all of its parts are organized and flow smoothly and logically from one idea to the next” (p. 48). According to this definition, they seem to define coherence at the message level, differentiating it from surface cohesion. They discuss three important features of coherence in terms of arranging sentences in a logical order, repetition of key words, and using transitional words. Although they emphasize the logic or flow of sentences, coherence is simplified to mere arrangements of sentences and using cohesive devices. In addition, only a limited range of cohesive devices is covered in their textbook. References and lexical cohesive devices such as using synonyms or related words in particular are not given proper weight in comparison to transitional expressions and conjunctions that are used to conjoin elements overtly. The lack of proper attention to lexical cohesive devices is rather unexpected, as lexical cohesion is one of the major cohesive devices, and effective writing is shown to display a variety of lexical means that go beyond simple reiteration (Connor, 1984; Halliday & Hasan, 1976).

According to Oshima and Hogue (2007), “a coherent paragraph flows smoothly from beginning to end” (p. 79). They suggest three main ways to make a text coherent: using nouns and pronouns consistently, using transitional signals (sentence connectors and conjunctions), and arranging ideas into logical order. Like the other two textbooks, Oshima and Hogue also highlight the use of transitional signals. They, however, do warn that overusing transition signals can be confusing. By logical arrangement of ideas, they mean dividing a topic into separate points and discussing them in turn. No special reference is made regarding lexical cohesion. The results are summarized in Table 1.

**TABLE 1**  
**Summary of the Three Textbooks**

	Fitzpatrick (2011)	Folse et al. (2010)	Oshima & Hogue (2007)
Coherence	Not discussed	Smooth flow and logical organization	Smooth flow of ideas
Sentence transitions/ Conjunctions	Emphasized in all units	Emphasized in all units	Emphasized in all units
Substitution/Ellipsis	Not discussed	Not discussed	Not discussed
Cohesion	Using consistent pronoun point of view Demonstratives are not discussed	Every pronoun refers to a specific noun Demonstratives are not discussed	Consistent use of pronouns Demonstratives are not discussed
Lexical cohesion	Lexical repetition and using related words	Repetition of key words	Not discussed

The five major cohesive devices, according to Halliday and Hasan (1976), are reference, substitution, ellipsis, conjunction, and lexical cohesion (see Section 2.1). As Table 1 shows, not all five are included and emphasized in the textbooks reviewed. The cohesive devices that all three textbooks include as important instructional targets are conjunctions and sentence transitions. As discussed in Section 2.2, nonnative speakers tend to overuse sentence transitions and conjunctions, and one can speculate that this tendency might be strengthened by unbalanced treatments of various cohesive devices, with too much emphasis on sentence transitions and conjunctions.

As for reference, an unbalanced treatment between personal and referential or demonstrative pronouns was found. The reviewed textbooks focus only on personal pronouns, but as pointed out in Hinkel (2001), L2 writers often overuse or misuse referential pronouns and demonstratives. This tendency has to do with underuse of lexical cohesive devices such as synonyms, paraphrasing, or collocations (Connor, 1984, Crossley & McNamara, 2009; Hinkel, 2011). Hinkel (2011) suggests that as L2 learners do not possess proper lexical means of expressing a previously mentioned referent in different words, they instead rely on referential pronouns or demonstratives, which apparently have simple forms. Yet the textbooks reviewed do not provide an adequate overview of a full range of lexical cohesion devices either. In the L1 writing literature, good writers are shown to employ a wider range of lexical cohesion beyond simple repetition (Witte & Faigley, 1981). L2 writing instruction needs to concentrate on the uses and functions of various lexical cohesive devices as well as on the proper use of referential and demonstrative pronouns.

No textbooks reviewed include substitution and ellipsis as their instructional targets.

Relatively little research has been conducted on substitution and ellipsis as cohesive devices. Future research is needed in this area, which in turn would shed light on how to treat the two in the teaching of cohesion. In addition, the textbooks tended to simplify the concept of coherence reducing it to mere organizational or connective flow. As sentence transitions or organizational patterns alone cannot make the text coherent, the teaching of coherence needs to emphasize how to enhance the clarity and logic of thoughts and ideas.

## 4. CORPUS ANALYSES

Among the different types of cohesive devices, as the textbook analyses suggested, sentence transitions were mainly dealt with in the three textbooks, but demonstratives were not. Based upon this result of the textbook analyses, this study explored whether learners' use of cohesive devices—sentence transitions and demonstratives—differ from native speakers' use by analyzing frequencies of cohesive devices in learner and native speakers' corpora. Also, this study analyzed the use of cohesive devices in writing by L2 proficiency level.

### 4.1. Data

This study used a learner corpus compiled in the College English Program at Seoul National University<sup>1</sup> for four years from 2006 to 2009. The corpus consists of Korean students' academic writing in English, which were required for the English courses entitled College English I and College English II, which offer instruction on how to write in an academic setting. The types of essays include narrative, descriptive, cause/effect, and persuasive writing, and topics vary accordingly (see Shin, 2011).

In analyzing the corpus based on proficiency level, Shin's (2011) three sub-corpora were used, which were formed according to TEPS scores ranging from 501 to 800 (three sub-corpora, hereafter designated as 500s, 600s, and 700s). According to the descriptions at the official TEPS website, TEPS scores ranging from 501 to 600 (i.e., the 500s corpus) indicate a mid-intermediate level of communicative competence, which are comparable to TOEIC scores ranging from 620 to 705 and TOEFL scores ranging from 180 to 207. TEPS scores ranging from 601 to 700 (i.e., the 600s corpus) indicate a high-intermediate level of communicative competence, which are comparable to TOEIC 710-780 and TOEFL 207-223. TEPS scores 701-800 (i.e., the 700s corpus) suggest that learners' proficiency has

---

<sup>1</sup> We are grateful to the College English Program at Seoul National University for allowing us to access this corpus.

reached an advanced level of communicative competence, which are comparable to TOEIC 710-855 and TOEFL 227-243. The numbers of essays and words contained in the three sub-corpora are shown in Table 2.

**TABLE 2**  
**Description of Sub-corpora**

Sub-corpus	500s	600s	700s
Number of essays	1,935	2,387	539
Number of words	435,255	542,212	129,845

## 4.2. Data Analysis Procedure

Based on the textbook analysis above, among the major cohesive devices, sentence transitions and demonstratives were used in the corpus analyses. In particular, the demonstrative pronouns *this*, *these*, *those* and the sentence transitions *however*, *therefore*, *in addition* were selected for the corpus analyses. The demonstrative pronoun *that* was excluded from this analysis because it was difficult to sort the demonstrative *that* from the data; *that* can be a complementizer or a relative pronoun in addition to a demonstrative pronoun. The sentence transitions (*however*, *therefore*, *in addition*) were chosen because they were the most frequent ones in the learner corpora among the representative sentence transitions (*however*, *therefore*, *in addition*, *on the other hand*, *in contrast*) which were found in common in all the three writing textbooks in Section 3.

In addition to the learner corpora, a large-scale native speaker corpus, the Corpus of Contemporary American English (COCA), was also consulted for usage of these target words. In particular, in order to compare the frequencies of target words in the learner corpora with those in the native speakers' written data, the academic sub-section of COCA was analyzed for the target words. The academic sub-section of COCA contains 82 million words from nearly 100 different peer-reviewed academic journals across diverse subjects.

Focusing on these target words, quantitative analyses were performed on the data. First, the occurrence of each target word was processed using the AntConc 3.2.1w concordancing program (2007). Second, the use of cohesive devices was examined by calculating their frequencies and chi-square statistics in the learner corpora in comparison to those in the COCA.

## 4.3. Results

Table 3 demonstrates raw and normalized (per 1,000,000 words) frequencies of the target words in the "academic" sub-section of COCA and the learner corpus.

**TABLE 3**  
**Frequencies of Target Words in the COCA and the Learner Corpus**

	Raw		Normalized	
	COCA	Learner	COCA	Learner
Demonstratives				
<i>this</i>	433,942	7,154	5,292	6,461
<i>these</i>	199,517	2,770	2,433	2,502
<i>those</i>	98,049	715	1,196	646
Total	731,508	10,639	8,921	9,608
Transitions				
<i>however</i>	76,869	1,522	937	1,375
<i>therefore</i>	25,057	1,145	281	1,034
<i>in addition</i>	22,915	557	279	503
Total	124,841	3,224	1,497	2,912
Grand Total	856,349	13,866	10,418	12,520

The normalized data can show direct comparisons between the frequencies of the target words in native speakers' writing and L2 learners' writing. The normalized data in Table 3 showed L2 learners' overall tendency to overuse cohesive devices, including both demonstrative pronouns and sentence transitions compared to native speakers.

The chi-square test on the normalized data showed that L2 learners' writing used more target cohesive devices than native speakers' academic texts ( $\chi^2 = 194.86, p < 0.01$ ). The normalized data in particular showed that the target sentence transitions were approximately twice as frequent in L2 learners' samples than in the native writers' samples ( $\chi^2 = 455.13, p < 0.01$ ), while the demonstrative pronouns were 1.07 times as frequent in the L2 corpus as in the native corpus ( $\chi^2 = 25.71, p < 0.01$ ).

Table 4 shows the overall frequencies of target words in the learner corpora by proficiency level. The normalized (per 1,000,000 words) frequencies are presented as well. The normalized data by proficiency level showed that the tendency to use more cohesive devices increased as learners' proficiency reached the high-intermediate level (TEPS 600s), whereas the use of the cohesive devices decreased as proficiency reached the advanced level (TEPS 700s). However, the frequency pattern of sentence transitions was different from the overall pattern that was mainly driven by pronouns. The normalized transitions data showed that the tendency to overuse transitions weakened as proficiency increased.

In order to check statistical significance between frequencies of each sub-corpus, pairwise chi-square tests were carried out on the normalized data, and the results are shown in Table 5.

**TABLE 4**  
**Overall Frequencies of Target Words in the Learner Corpora**

	Raw			Normalized		
	500s	600s	700s	500s	600s	700s
<b>Demonstratives</b>						
<i>this</i>	2,666	3,629	859	6,125	6,693	6,616
<i>these</i>	1,048	1,422	300	2,408	2,623	2,310
<i>those</i>	258	354	103	593	653	793
Total	3,972	5,405	1,262	9,126	9,968	9,719
<b>Transitions</b>						
<i>however</i>	624	717	181	1,434	1,322	1,394
<i>therefore</i>	461	594	90	1,059	1,096	693
<i>in addition</i>	243	269	45	558	496	347
Total	1,328	1,580	316	3,051	2,914	2,434
Total	5,300	6,985	1,578	12,177	12,882	12,153

**TABLE 5**  
**Results of Chi-square Test Comparing Three Sub-corpora ( $\chi^2$ )**

	Overall	Pronouns	Transitions
700s/600s	21.50**	3.18 <sup>†</sup>	43.20**
600s/500s	20.09**	37.49**	3.16 <sup>†</sup>
700s/500s	0.02	18.84**	69.60**

Note. \*\*  $p < 0.01$ ), \*  $p < 0.05$ ), †  $p < 0.10$

The overall tendency to rely on cohesive devices in the 600s sub-corpus was significantly different from that in the 500s and 700s sub-corpora; the 600s sub-corpus contained more cohesive devices than the 500s and 700s. In addition, for types of cohesive devices, the results showed that the 600s and 700s sub-corpora contained significantly more demonstrative pronouns than the 500s sub-corpus. The tendency to use sentence transitions, however, decreased as proficiency increased. The number of sentence transitions in the 700s sub-corpus was significantly less than those in the 500s and 600s. However, one should note that the advanced learners still overused transitions 1.63 times more than native speakers (2,434 vs. 1,497;  $\chi^2 = 223.78$ ,  $p < 0.01$ ).

## 5. DISCUSSION AND CONCLUSION

This study examined cohesive devices in English writing textbooks and learners' writings through text and corpus analyses. The results of the text analyses showed that three writing textbooks did not cover a full range of cohesive devices; instead, they focused

on sentence transitions and conjunctions, which L2 writers often overuse or misuse. Of course, sentence transitions and conjunctions are one of the most important means to connect and organize linguistic units and ideas, and it is not surprising that textbooks include them as important instructional objectives. Also, many L2 writing textbooks, including the ones reviewed in this study, are organized based on paragraph or essay types such as cause-effect, contrast-comparison, or argumentation, which may make it feasible to include different sentence transitions associated with a specific writing type. The three textbooks analyzed in this study, on the other hand, did not fully cover demonstratives and lexical cohesive devices such as synonyms, paraphrasing, or collocations. Although individual instructors may supplement their textbooks and the contents of the textbooks may not exactly mirror pedagogical practice, the review suggests that there is a room for improvement in how English writing textbooks approach coherence and cohesion. Future research on a wider range of textbooks and actual instruction practice can provide insights into how L2 coherence and cohesion are pedagogically approached.

The corpus analyses showed that L2 writers used more target cohesive devices in their writings than native writers. However, the tendency to overuse sentence transitions grew weaker as the learners' proficiency increased, but the 700s group still used 1.63 times as many sentence transitions as the native group. The over-application of sentence transitions can lead to texts that are difficult to follow and that sound discontinuous. Also, overuse of transitions often involves illogical or inappropriate usages. For example, in the following excerpt (1) from an essay arguing for the safety of genetically modified agricultural products, the use of additive sentence transitions *moreover* and *in addition* does not seem to be warranted.

(1) In Netherlands, the transgenic potatoes with improved storage have already been developed, and they are expected to create more earn for the potato farmers. Moreover, the world can provide more food of various kinds to the poor because GM plants' high storability will create lots of surplus produce. In addition, surplus produce will cause the price cutting of the crops, and price cutting means people can access to food more easily than before. (TEPS 600s)

Additive sentence transitions such as *moreover* and *in addition* supply information building on previous material. Importantly, a new point and the preceding point should be "parallel" (Cowan, 2008). In the example (1) above, the writer is arguing that genetically modified products have advantages in terms of longer storage periods. At the idea level, the three benefits—more earning for farmers, provision of food to the poor, and easy access to food—may seem parallel, but as the three sentences in the excerpt have different grammatical subjects, the use of additive sentence transitions without surface parallelism

does not help creating cohesive ties among the ideas.

Similarly, in (2), the sentence transition *therefore* is employed in an inappropriate way.

(2) Abuse of SNSs can negatively affect teenagers' everyday life. However, there is no government's counterplan about bad impacts of teenagers' SNS use. Therefore, parents should regulate the use of SNS of their teenage kids. (TEPS 700s)

In this excerpt, the lack of government policies regulating teenagers' use of social network services does not provide a clear logical support for why parents should take actions. As these examples show, L2 learners' tendency to overuse sentence transitions is not just a matter of increased frequency; the problem is that oftentimes learners use sentence transitions somewhat loosely in contexts where they are not appropriate.

Learners also tended to overuse demonstratives as they reached a certain level (around TEPS 600-700) as the following example (3) shows:

(3) Nursing reality is usually covered under the surface, so people don't know well about the reason why this is problems the government should resolve. ... People should also know the reason why the nurse-to-patient ratio is an important controversial issue. ... Nurses should pay attention to patients' every conditions and medical histories from A to Z and continuously take care of them. In this reality, how do they deal with ten to fifty people in an hour? This is a disaster not only to nurses, but also to patients and their families because they aren't capable of being received right treatment perfectly. Again, the government should take measures to solve staffing problems, because this is directly related to patient safety. (TEPS 700s)

In (3), the writer uses the demonstrative pronoun *this* quite frequently, but not all of these uses are appropriate. For example, the writer uses demonstratives without a specific referent (e.g., *This is a disaster*), or without collocationally fitting or grammatically congruent referents (e.g., *this is problems*, *this is directly related*), making the text confusing. In texts written by native speakers, on the other hand, demonstratives have clearly identifiable antecedents rather than referring to implied or non-explicit ideas (Halliday & Hasan, 1976; Hinkel, 2001; Kim, 2012). As shown in (4), an excerpt from an essay written by a native speaker which won a prize in the 2004 Humanities Essay Contest, demonstratives substitute identifiable referents (e.g., *This theory*, *this exposure*) often involving paraphrasing (e.g., *this decentralized medium*, *these protectionist policies*).

(4) The Internet acts as another vehicle for the worldwide propagation of American influence. Interestingly, some commentators cite the new "information economy" as

proof that American cultural imperialism is in decline. They argue that the global accessibility of this decentralized medium has decreased the relevance of the “core and periphery” theory of global influence. This theory describes an inherent imbalance in the primarily outward flow of information and influence from the stronger, more powerful “core” nations such as the United States. ... Thus, the Internet can dramatically increase exposure to American culture for those who desire it. Fear of the cultural upheaval that could result from this exposure to new information has driven governments in communist China and Cuba to strictly monitor and regulate their citizens’ access to websites (these protectionist policies aren’t totally effective, however, because they are difficult to implement and maintain). (Native speaker)

As the comparison suggests, L2 writers oftentimes use demonstratives somewhat loosely. Apparent simplicity of the form, lack of proper attention in textbooks and instruction, and learners’ limited linguistic resources to paraphrase previously mentioned information might all contribute to the frequent yet inappropriate application of demonstratives. For the effects of learners’ proficiency, the corpus results did not yield any meaningful differences between the 700s group and the 600s group in their use of demonstrative pronouns. Analyzing actual samples from different proficiency groups by examining appropriate and inappropriate uses of demonstrative pronouns will help to clarify the developmental trajectory of demonstratives.

Although a direct comparison between the academic sub-section of COCA and L2 academic writings might not be appropriate due to different genres of the corpora (academic journals vs. academic essays) and no direct cause and effect relationship between the results of text analyses and corpus analyses can be established, these results can yield important implications for L2 writing instruction. L2 writers’ frequent use of sentence transitions does not seem to be unrelated to the textbooks and instruction in L2 writing classes. Rather than blind emphasis on sentence transitional expressions, L2 writing instruction will need to deal with semantic and syntactic functions and usages of different transitional expressions and demonstrative pronouns (Granger & Tyson, 1996). Also, it can be predicted that instruction in a variety of cohesive devices including synonyms, paraphrasing, or collocations will help equip learners with varied means of cohesive devices. Qualitative comparisons between native and learner samples will help better understand the causes for the overuse and misuse of certain cohesive devices.

## REFERENCES

- Ammon, U. (2001). *The dominance of English as a language of science*. Berlin/New York: Mouton Gruyter.
- Aston, G. (1997). Small and large corpora in language learning. In B. Lewandowska-Tomaszczyk & J. P. Melia (Eds.), *Practical applications in language corpora* (pp. 51-62). Lodz, Poland: Lodz University Press.
- Bolton, K., Gerald, N., & Joseph, H. (2003). A corpus-based study of connectors in student writing. *International Journal of Corpus Linguistics*, 7(2), 165-182.
- Campbell, K. S. (1995). *Coherence, continuity, and cohesion*. Hillsdale, NJ: Lawrence Erlbaum Associates.
- Carrel, P. L. (1982). Cohesion is not coherence. *TESOL Quarterly*, 16(4), 479-488.
- Choi, Y. (1988). Text structure of Korean speakers' argumentative essays in English. *World Englishes*, 7(2), 129-142.
- Connor, U. (1984). A study of cohesion and coherence in ESL students' writing. *Papers in Linguistics: International Journal of Human Communication*, 17(3), 301-316.
- Cowan, R. (2008). *The teacher's grammar of English: A coursebook and reference guide*. Cambridge: Cambridge University Press.
- Crossley, S. A., & McNamara, D. (2009). Computational assessment of lexical differences in L1 and L2 writing. *Journal of Second Language Writing*, 18(2), 119-135.
- Ellis, R. (1994). *The study of second language acquisition*. New York: Oxford University Press.
- Fitzpatrick, M. (2011). *Engaging writing 1: Essential skills for academic writing*. New York: Pearson Longman.
- Folse, K. S., Solomon, E. V., & Clabeaux, D. (2010). *Great writing 3: From great paragraphs to great essays* (2nd ed.). Boston, MA: Heinle.
- Granger, S. (Ed.). (1998). *Learner English on computer*. New York: Longman.
- Granger, S., & Tyson, S. (1996). Connector usage in the English essay writing of native and nonnative EFL speakers of English. *World Englishes*, 15(1), 17-27.
- Halliday, M. A. K., & Hasan, R. (1976). *Cohesion in English*. London: Longman.
- Hinkel, E. (2001). Matters of cohesion in L2 academic texts. *Applied Language Learning*, 12(2), 111-132.
- Hinkel, E. (2004). *Teaching academic ESL writing: Practical techniques in vocabulary and grammar*. Mahwah, NJ: Lawrence Erlbaum Associates.
- Hinkel, E. (2011). What research on second language writing tells us and what it doesn't. In E. Hinkel (Ed.), *Handbook of research in second language teaching and learning* (pp. 523-538). New York/London: Routledge.
- Hyland, K. (2006). *English for academic purposes*. London/New York: Routledge.

- Johns, T. (1991). Should you be persuaded—two samples of data-driven learning materials. *English Language Research Journal*, 4, 1-16.
- Kang, J. Y. (2009). Referencing in a second language: Korean EFL learners' cohesive use of references in written narrative discourse. *Discourse Processes*, 46(5), 439-466.
- Kim, Y.-J. (2012). An analysis of referential use in Korean EFL learners' argumentative essays. *English Teaching*, 67(3), 181-204.
- McCarthy, M., & O'Keeffe, A. (2010). Historical perspective: what are corpora and how have they evolved? In A. O'Keeffe & M. McCarthy (Eds.), *The Routledge handbook of corpus linguistics* (pp. 3-13). New York: Routledge.
- Oh, S.-Y. (2007). A corpus-based study of epistemic modality in Korean college students' writings in English. *English Teaching*, 62(2), 147-175.
- Oh, S.-Y. (2009). Korean college students' use of English demonstratives in argumentative essays. *English Teaching*, 64(1), 51-78.
- Oshima, A., & Hogue, A. (2007). *Introduction to academic writing* (3rd ed.). New York: Pearson Longman.
- Park, Y.-Y. (2013). Korean college EFL students' use of contrastive conjunctions in argumentative writing. *English Teaching*, 68(2), 55-77.
- Shin, J.-A. (2011). Overpassivization errors in Korean college students' English writings. *Korean Journal of Applied Linguistics*, 27(3), 255-273.
- Witte, S. P., & Faigley, L. (1981). Coherence, cohesion and writing quality. *College Composition and Communication*, 32(2), 189-203.
- Yoon, J.-W., & Yoo, I. W. (2011). An error analysis of English conjunctive adjuncts in Korean college students' writing. *English Teaching*, 66(1), 225-244.

Applicable levels: College

Hee Youn Cho  
College English Program  
Faculty of Liberal Education  
Seoul National University  
1 Gwanak-ro, Gwanak-gu  
Seoul 151-742, Korea  
Email: heecho@snu.ac.kr

Jeong-Ah Shin  
Department of English Language and Literature  
Kwangwoon University  
20 Kwangwoon-ro, Nowon-gu,

Seoul 139-701, Korea  
Email: jashin@kw.ac.kr

Received in December 2013  
Reviewed in January 2014  
Revised version received in February 2014

