Korean ESL Learners' Use
of Connectors in English Academic Writing

Kent Lee

(Korea University)

Some previous studies have examined the use of connectors by second language writers of English, particularly studies comparing connector usage of L2 and native English writers. This study does so with corpora of written essays, comparing the frequency of connector usage between native English writers from the U.S. and U.K., and English essays by Korean ESL learners in the US. The statistical comparisons used here also offer some refinements over previous corpus studies. The Korean writers overused and underused various connectors compared to the native writers, specifically in their use of contrastive markers, topic transitionals, *there is* and *there are,* markers for enumerating sequences, and exemplifiers. They also certain expressions used incorrectly as connectors due to apparent transfer errors. These data indicate a number of issues for ESL/EFL writing teachers to address, and some suggestions for teaching are discussed below.

I. INTRODUCTION

Discourse connectives impart logical and textual coherence to texts. Reading experiments show that discourse connectives facilitate and readers' comprehension of texts, in that they provide explicit cues about the logical relationships among referents, propositions, and clauses, and thus help readers to construct online mental representations of the meaning of a text (Degand, Lefèvre, & Bestgen, 1999). Subordinating conjunctions also facilitate reading times and coherence by marking information as relatively secondary, so that readers can focus more on the content of main clauses (Traxler & Gernsbacher, 1995). However, excessive use of connectors may not be felicitous, as some studies suggest that explicit markers can interfere with readers' text comprehension (Millis, Graesser, & Haberlandt, 1993). To what degree such relationships need to be made explicit depend on the genre or text type, and the readers' level of expertise and background knowledge (McNamara & Kintsch, 1996).

In recent years, applied linguists have begun to conduct corpus-based studies of second language (L2) writing in English, particularly by comparing native English writers and L2 in their use of connectors (i.e., connectives, transitionals). Some studies have examined ESL/EFL writers from Western language backgrounds, such as Granger and Tyson's (1996) study of

English connective use by L2 writers, compared to native writer (L1) patterns. Their study revealed patterns of overuse, underuse, and sometimes infelicitous use of particular transitional expressions by L2 writers of French and other European backgrounds. A number of other studies have found lexical, pragmatic and syntactic problems with ESL/EFL writing by learners of different language backgrounds, including a few studies of Korean ESL/EFL writers. Many of these studies have been descriptive, while a few comprehensive quantitative studies have been reported. Certain gaps exist in our quantitative understanding of Koreans' ESL or EFL writing patterns, which are examined and addressed in the study below. After surveying previous studies, identifying outstanding questions, and reporting the results of this current study, some implications for the teaching of academic English writing will be sketched out below.

II. LITERATURE REVIEW

In recent years, corpus based studies have shown that connector use can be problematic for second language writers, as their writing shows different patterns and frequencies of usage in logical and temporal connectors. Granger and Tyson (1996) was perhaps the first major comparative corpus-based study of second language writing. This study particularly compared a corpus of academic English essays by Francophone college students, and a corpus of L1 academic essays by U.K. and American college students. They reported patterns of overuse, underuse and incorrect usage of English connectives, particularly, overuse of the contrastive marker *to the contrary* and the additive marker *moreover*. These were explained as transfer errors; for example, the French *au contraire* is often translated as *to the contrary*, but its semantic scope and usage differs from the English counterpart, hence their overuse and pragmatically infelicitous use of this phrase. They also reported infelicitous use of *i.e.,* colloquial wording, and colloquial transitionals such as *first of all.* These L1 and L2 writer data sets comprise the International Corpus of Learner English (ICLE), as described subsequently in Granger (2002).

A number of similar studies have investigated similar ESL/EFL issues, mostly involving learners from European language backgrounds; for example: tense use among French writers (Granger, 1999); register appropriateness and collocational usage among Europeans (Biber & Conrad, 2001); modal overuse among Swedish writers (Aijmer, 2002); overuse of high frequency verbs among various Europeans (Ringbom, 1999); and acquisition of the English verbal system by French and writers Dutch (Housen, 2002). Not surprisingly, East Asian students exhibit somewhat similar problems; for example, some descriptive studies have shown that Chinese writers overuse some English connectives such as *so, also, besides, therefore,* and *moreover,* and the concessives *but* and *however* (Milton & Tsang, 1993; Yang & Sun, 2012; Zhang, 2000). A quantitative corpus study of advanced Japanese EFL learners showed that

Japanese overuse some connectors such as *for example, of course,* and *first,* while underusing others like *then, yet,* and *instead* (Narita, Sato, & Sugiura, 2004). One study of Korean writers' verb usage (Cowan, Choi, & Kim, 2003) compared the ICLE L1 subset with their own corpus of Korean ESL writers. They reported transfer errors, particularly with incorrect passivization of unaccusative intransitive verbs (change of state verbs and verbs of existence), e.g., *was occurred, *was happened,* and *was existed*; such patterns are due to pragmatic and syntactic differences between the English and Korean passive voice.

Relatively few studies of Koreans' connector usage in ESL/EFL writing have been published. One study (Kang, 2005) found that in English narrative writing, Koreans rely heavily on pronominal and demonstrative pronouns as cohesive devices. This study also found that Koreans use significantly more conjunctions in narratives than native English writers, in frequency counts of adversative (contrastive), additive, causative, temporal, and other conjunctions. In English oral narratives, Koreans also have problems in native-like cohesion due to overuse of nominals (Kang, 2004). Another corpus study (Yoon & Yoo, 2011) found that Korean writers overuse sentence-initial conjunctions, even in connecting sentences that are overly short for such connectors; students made more errors with sentence fragments and run-on sentences; and students make many punctuation errors with conjunctive adverbs (e.g., *therefore, however*).

Most of these previous studies have used at most descriptive statistics (counts and percentages), while only a few studies have done full quantitative analyses of connector usage with inferential statistics, namely, Kang (2004), Kang (2005), and Narito et al. (2004); these studies of Korean and Japanese writers used ANOVAs and MANOVAs. Previous studies, however, have not examined L2 writers at different proficiency or course levels, but have mainly examined college level writers in general. A quantitative comparison of different L2 levels with L1 writers would be helpful for detecting patterns of change or improvement in connector usage from lower to higher levels. Also, ANOVAs and MANOVAs lack statistical power in detecting lexical frequency patterns, since word frequencies do not follow a normal distribution, but rather, a logarithmic distribution (Baayen & Lieber, 1997). Other statistical techniques would be more suitable for word count data, such as the loglinear models used in this study. This study is thus designed to address these gaps in the literature regarding Korean L2 writers of English.

In light of past research, one would naturally expect Korean ESL/EFL learners to show patterns of overuse, underuse, and infelicitous usage. This leads to the following questions: (1) what non-native-like patterns exist across different types of connectors; (2) what infelicitous expressions are Koreans likely to use; and (3) are Korean ESL/EFL writers likely to show improvement at higher proficiency levels, or do these errors persist into higher levels? The next section explains the study design used to address these questions.

III. STUDY DESIGN

To better understand the academic writing patterns of Korean L2 writers, a comparative study was carried out, comparing connector use between L1 and L2 writers. The data sets used in this study are (1) the native English writer subset of Granger's ICLE collection (Granger, 2003), and (2) the Korean L2 Corpus from Cowan et al.'s (2003) study. The following types of connectors were examined, based on those examined in past studies, the data available in the L2 corpus, and commonly used infelicitous expression that emerged upon examination of the corpus data: (1) so-called adversative and concessive markers (e.g., *but, although*), subsumed here under the more general rubric of contrastive markers; (2) exemplifiers such as *e.g.,* including infelicitous terms like *in case of*; (3) the emphatic adverb or focus marker *especially*; (4) enumerative markers (*first, second*, etc.); (5) *there is* and *there are*; and (6) topic shift markers for topic shifts.[1]

1. Materials

The L1 corpus is from the L1 writer subset of the International Corpus of Learner English (ICLE), consisting of college essays of native speakers from the US and UK compiled over several years (Granger & Tyson, 1996; see also Granger, 2000; 2003). The Korean L2 Corpus from Cowan et al. (2003) consists of data from normal writing assignments from foreign students in writing courses at a non-credit intensive English program (IEP), as well as writing assignments from foreign undergraduate and graduate students taking required remedial ESL writing courses. The writing samples were collected from students at the University of Illinois (Urbana-Champaign) in the US, who gave permission for their writings to be collected and anonymously. The samples were collected over a couple of years. The corpus includes the following course levels: (1) upper levels of IEP writing courses, collapsed here into one level[2]; (2) three undergraduate writing courses, collapsed into one; (3) ESL 400, a lower level ESL writing course; (4) ESL 401, an intermediate graduate ESL writing course; (5) ESL 402, an upper level graduate level ESL elective for thesis writing; and (6) ESL 405, an ESL writing and oral communication course for graduate students in business related fields. The essay were

---

[1] For these classifications of transitionals, see standard references such as Hinkel (2002), Hinkel (2004); for focus markers, see König (1991); for topic shift markers, see Schiffrin (1987) and Brinton (1996).

[2] IEP students are non-credit students (typically lacking previous communicative Enlgish experience, and having TOEFL scores too low for university entrance); they are placed into levels via in-house placement exams.The placement exams for IEP and regular students include writing components, and were developed by testing experts in the university's TESOL and Educational Psychology departments. Only students in the higher IEP levels took writing courses; thus, only more advanced IEP levels are represented here. Due to the relative paucity and shortness of IEP essays, they were collapsed together here;.their language ability is demonstrably lower than the undergradaute and graduate students (Cowan et al., 2003). Essays from various undergraduate courses and grade levels were also all grouped together for a larger sample size.

collected and transcribed anonymously, so the number of writers or number of essays per writer is unknown.

Undergraduate and graduate international students are placed into these levels via in-house language placement tests, and if required to take an ESL writing course, they do so in their first semester. Thus, almost all of them were new foreign students at the time, who had just arrived from Korea to begin their studies in the U.S. (except for some of the undergraduate students who may have been in the U.S. previously). These placement levels serve then as a proxy for language proficiency levels. Word counts for these levels are summarized below, along with those for the L1 corpus of native English writers (Granger, 2002). Although the L1 corpus has fewer essays from fewer subjects, it contains considerably longer writing samples (see Table 1).

TABLE 1
Summary of L1 and L2 Corpora

corpus	level	essays	word count
Korean L2 Corpus	IEP	56	11,822
	undergraduate (UG)	110	60,236
	ESL 400	80	29,134
	ESL 401	192	115,988
	ESL 402	30	17,686
	ESL 405	22	6690
	total	490	241,556
L1 Corpus	undergraduate	376	95,695

2. Method

Categorical data analysis techniques were used, namely, chi-square tests based on loglinear regression analysis (Agresti, 1996), which are more appropriately for logarithmically and otherwise non-normally distributed data such as lexical frequencies (Baayen & Lieber, 1997). Loglinear tests are based on logarithmic functions, and readily allow for comparisons of such data[3]. Frequency counts for discourse connectives were counted (using Linux fgrep search tools) and compared, first between the Korean L2 Corpus and the L1 Corpus to look for significant differences between native and non-native speakers in usage of particular lexical items. When possible, connective usage was compared between the different course levels of the Korean L2 corpus to look for developmental changes in usage patterns. That is, at higher levels students

---

[3] That is, the regression equation is based on a logarithmic function, allowing for analysis of non-normally distributed data. This also allows for regressing on nominal variables, including dependent or outcome variables. In these data, a writer's choice of connective is not numerical or ordinal, put a purely nominal outcome variable, for which loglinear methods are appropriate. See standard references such as Agresti (1996) for more on loglineary regression, and Baayen and Lieber (1997) on its use in linguistics studies.

should commit fewer errors, and should be more able to avoid overuse and underuse of particular markers which students in lower levels overuse or underuse. Because of the smaller word counts in certain levels, the data in some cells were too few for statistical analysis.

IV. RESULTS AND DISCUSSION

Results of the comparisons are reported below for contrastives, exemplifiers, *especially*, enumeratives, *there is/are*, and topic shift markers, including some awkward transitional expressions used within these categories.

1. Contrastive Markers

The chief contrastive markers in English include *but, yet, however,* and *although*. Their Korean equivalents are the conjunctive adverbs *kurena, kurehciman, kurendei* (*kundei*) *haciman*, and the enclitic *-ciman*, which attaches to verbs. One difficulty for Koreans is the apparent lack of one-to-one correspondence between these Korean and English connectives. Other contrastive markers like *while* and *whereas* did not occur in the L2 corpus, or were far too few for statistical analysis. Non-contrastive uses of *yet* as an aspectual marker (König, 1991), i.e., a temporal adverbial as in *not yet*, were excluded.

Significant differences were found in the distribution of the contrastive markers between Korean and English native writers. An overall χ^2 of 73.74, p<.0001 was found between both groups, indicating that Korean L2 students tended to overuse contrastive markers in general, especially *but*, but lacked stylistic or rhetorical variation in that they greatly underused some markers such as *yet*. When compared to the total word count as a rate of occurrence in each corpus, Korean L2 students used contrastive markers 7.2% of the time (1740 tokens), while native English writers used them 5.3% of the time (483 tokens). The Korean students were thus explicitly marking contrasts more often than native writers, while native writers more often made implicit statements of contrasts. Overall, the L2 writers used *but* and *though* significantly more than the L1 writers, and the L2 writers used the other contrastive markers *although, however* and *yet* significantly less than the L1 writers (*but:* χ^2=33.7, p<.0001; *although:* χ^2=10.3, p<.067, marginal effect; *though:* χ^2=34.4, p<.0001; *however:* χ^2=15.2, p=.0096; *yet:* too few L2 tokens for analysis). The one marker that seems less problematic is *though*, which is more common than *although* in informal English, and students may thus be more familiar with it, and it is syntactically similar to the Korean subordinator *-ciman*.

Frequencies of contrastive markers at different levels were compared to determine whether higher level students showed more native-like patterns. These tests showed that the

overuse of *but* did not attenuate with higher ESL levels; overuse of *though* attenuated slightly but was still overused at the higher levels; underuse of *although* and *however* did not improve significantly at higher levels; and tokens of *yet* were too few for comparison, also indicating that *yet* was greatly underused. Frequencies are reported in Table 2, with percentages for the number of contrastive markers per group, e.g., *but* constitutes 52.5% of contrastive markers used by the L2 writers, and 50.3% of those used by the L1 writers. A more detailed analysis of these markers will be presented in a future study.

TABLE 2
Contrastive Markers by Level in the KL2 Corpus, with Total Word Counts per Level

conj.	IEP	UG	400	401	402	405	L2 total		L1 total	
but	64	205	122	463	32	29	915	52.50%	255	50.30%
though	17	64	26	65	8	2	182	10.44%	21	4.14%
although	6	34	13	48	11	1	113	6.48%	50	9.86%
however	19	147	67	240	40	11	524	30.06%	146	28.80%
yet	0	5	1	0	0	0	6	0.005%	11	4.73%
TOTAL	106	455	229	816	91	43	1740	100%	483	100%

A few minor errors were found with several contrastive markers, in that students tended to mark them with commas, which is not normally be done in academic writing (except occasionally for emphasis); see Table 3 for examples. Three comma errors with *even though* were found, one each in the undergraduate, 400, and 401 levels. One example of contrastive *yet* was found sentence-finally. Most comma errors were with sentence initial *but* (70 tokens). While *but* plus comma is not an egregious error, it leads to a colloquial tone, as it is relatively rare in formal writing, and reflects a colloquial use of *but* followed by a pause for stronger contrast or emphasis, e.g., *But, I really don't know* (Schiffrin, 1987, p. 162). The marker *but* has multiple analogues in Korean, often followed by an intonational juncture before the rest of the clause, which may lead Koreans to assume that *but* is also pause-marked, and thus, takes a comma in writing. Comma errors were as follows: IEP, 6 tokens; UG, 4 tokens; ESL 400, 7 tokens; ESL 401, 47 tokens (due to this group's larger sample size); ESL 402, 4 tokens; and ESL 405, 2 tokens. The L2 writers, overall and in each ESL level, used commas with *but* significantly more than the L1 writers, who did not use *but* with commas (χ^2=265.6, p>.001). This informal comma usage did not seem to attenuate at higher levels, indicating that this is a persistent error, such that even some Korean graduate students were not aware of its usage.

TABLE 3
Comma Errors with Contrastive Markers

conj.	example
but	Murray Weidenbaum argues that the much money the government spends is never sufficient for eliminating poverty, and it makes the situation of the poor worse, not better. <u>But,</u> statistics show the greatest escape rates from poverty by welfare benefits. (ESL 401 sample)
even though	In fact, we are exposed to unexpected hazards, as our society become more complicated. Even so, can we say that there are more hazardous than before? Maybe almost all people say "absolutely no." <u>Even though,</u> why do many people feel fear, furthermore phobia through their life and strive to get out of them in vain. (ESL 400 sample)
yet	No one can tell whether ADHD is personality characteristics or pathology, <u>yet</u>. Thus, making a diagnosis of ADHD and prescribing methylphenidate like Ritalin should be considered prudently. (UG sample)

2. Exemplifiers

Another category that was noticeably problematic for Korean students is a group of connectives that are designated here as exemplifiers. This group consists of items such as *especially, such as, specifically, for example,* and *e.g.,* which are used to indicate elaboration and exemplification. They mark transitions from a more general point to a more specific supporting point or argument, or transitions to a specific supporting example. Since they also mark transitions to new discourse topics, they are functionally related to the discourse topic transitionals discussed in the next section that mark shifts to previous or related topics. Exemplifiers like *for example, in the case of* likewise indicate a shift to a specific items cited as evidence for the preceding point. The Korean writers also used *for example* more than other functionally similar exemplifiers, and rarely used the more formal academic forms like *e.g.* and *i.e.* as shown in Table 4.

TABLE 4
Exemplifier Counts

	for example	e.g.	i.e.
IEP	15	0	0
UG	43	0	0
400	132	0	0
401	87	0	1

402	15	5	2
405	1	1	1
Korean L2 total	293	6	4
L1	44	6	9

Two other types of exemplification markers appeared in the corpus that were used excessively and awkwardly. These are phrases used sentence initially for introducing supporting examples. The first is the *case of* marker, in the form *in [the] case of, in my case*, which is influenced by the original Korean usage of *kyĕngwu* 'case' and *kyengwu-e nun* 'in case of, in X case.' Some examples from the L2 corpus are shown below:

> According to Lieberman..., "the sounds that Neanderthals could not have made because of their supralaryngeal vocal tract anatomy have functional properties that enhance speech perception." In case of tongue, the Neanderthal tongue is contained within the oral cavity (ESL 401 sample).

> In conjunction with the language barrier, international students have a big trouble in understanding classes. In my case, the stresses that I felt was extreme when classes began (ESL 400 sample).

In Korean, this is a more common transitional expression to introduce an item as a specific example to a more general claim. Phrases like *kyengwu* are grammaticalized as topic transitional markers with the addition of the topicalizing particle *nun*, which effectively marks it as a sentence topic, although it is new to the discourse, e.g., *Hankuk-ui kyengwu-ey-nun,* which translates literally as *in the case of Korea*; or *che-ui kyengwu-ey-nun, in my case.* In this respect, it is morphosyntactically and functionally similar to the topic shift markers discussed in a following section.

While native English writers in the L1 used no such transitional expressions, the *case of* form appeared 26 times in the Korean L2 Corpus, and *in my case* appeared 13 times. The breakdown by levels is given in Table 5. Since the L1 writers did not use such transitional expressions, a statistical comparison would be superfluous.

TABLE 5
Frequencies of *case of* Expressions

level	case of	in/for my case	total
IEP	2	3	5
UG	1	2	3
400	3	5	8
401	19	3	22

402	0	0	0
405	1	0	1
total L2	26	13	35

The second pragmatically anomalous marker is a Korean marker directly translated into English in the form *as evidence*[*s*]. This is a direct translation of the Korean transitional *cungge*, which has no suitable English equivalent. It can appear clause-initially as *cungge-ro* "*as evidence*" to introduce a supporting fact or example in argumentation. Three occurrences of this marker were found, such as this example.

> Generally many people believe that Countries with national health insurance hold down costs by operating more efficient health care systems. As evidences, many defenders of national health insurance give the fact that the Socialized health care reduces the cost in other countries by managing efficiently (ESL 401 sample).

3. *Especially*

The adverb *especially* in formal academic writing is usually an intrasentential modifier of a verb or adjective, e.g., *She especially likes persimmons*. As such, it functions as an emphatic and restrictive marker (König, 1991), in that it selects and emphasizes one possibility from among a set of other hypothetical possibilities (e.g., primarily liking persimmons, versus other known or implied possibilities). As a clause-initial or sentence-initial sentential adverb, it takes on an exemplifying or evidential function in addition to its restrictive function, and as a clause-initial adverb, it puts emphasis on the entire clause. Thus, it seems out of place in a sentence like *She spends a lot of time in the orchard; especially, she likes persimmons,* but is more felicitous in a sentence like *She spends a lot of time in the orchard; especially in the spring, she gives her attention to the persimmons*, where this adverb modifies a local phrase rather than the whole clause. This use is much generally colloquial, and much less common in formal academic English writing.

In the Korean L2 corpus, a number of examples of the colloquial sentential adverb *especially* appear. While this is not an error per se, it can lead to an undesired colloquial tone. The Korean writers showed no unusual usage of *in particular* or *particularly*, which are felicitous sentence-initially. An example appears below, and Table 6 compares counts for *especially* as initial sentential adverbs and as a regular intra-sentential manner adverb.

> TQM philosophy is crucial to enhance internal value chains and external satisfactions to business management. As a result, it is also needed to socialization of human activities. Especially, business activities attempt to efficiency and satisfaction in the marketing. (ESL 401

—

sample).

TABLE 6
Ratios of *especially* as Sentential and Manner Adverb

corpus	sentential adv.	manner adv.	total
IEP	3	2	5
UG	10	24	34
400	5	18	23
401	25	46	71
402	1	15	16
405	1	4	5
Korean L2 total	44	112	156
L1	0	31	31

Since the L1 writers did not use *especially* as a sentence-initial adverb at all, a statistical comparison between the L1 and L2 corpora is essentially superfluous. The Korean L2 writers used it excessively as a sentential adverb, while the native English writers only used it intrasententially as a manner adverb. One would expect to see improvement at higher proficiency levels, with less frequent usage of the initial sentential adverb less compared to the intrasentential modifier adverb. An overall trend test was marginal (χ^2=12.0, p=.062). However, the data for the top two ESL levels seem to indicate some improvement at higher levels. Also, the lower IEP and ESL 400 levels fared significantly worse (χ^2=4.32, p<.038) than the other ESL groups. This would suggest that students in those levels significantly overuse the sentential adverb, and that those in the upper levels seem to show some improvement.

This excessive usage of *especially* is influenced by the distribution of its Korean counterparts, the marker *thukhi* or *thukbyelhi*. These Korean markers can be used more freely as an initial sentence adverb, and with a wider range of meanings, compared to its English counterpart, e.g., *thukhi cu-ui hada* 'pay special attention to' and *thukhi cicek hada* 'expressly point out.' As a result, distinction between these two functions may be less clear to Korean L2 students, leading to their overuse of this term.

4. Enumerative Markers

A disproportionately large number of enumeratives were found in the L2 corpus, particularly ordinals such as *first, second, finally,* and their British analogues *firstly, secondly,* etc. These were used by the Koreans much more than native English writers, as seen in Table 7. While this is not a serious error, they are not used as often in formal and academic English as the L2

students have used them. This can seem slightly awkward pragmatically and detract from the tone they wish to convey, in that this overuse can make their writing seem more mechanical to readers.

TABLE 7
Number of Tokens of Enumerative Markers

	IEP	UG	400	401	402	405	L2 total	L1
first	10	25	10	24	16	4	89	
firstly		2	2	4		1	9	15
first of all	3	12	13	22	7	2	59	1
second	10	16	12	21	7	4	70	
secondly		12	6	18	2	2	40	6
third	5	4	4	8	2	1	24	
thirdly		1		7	2		10	1
fourth	1	1		1			3	
fourthly		1		1			2	
fifth	1						1	
next	5		4		5		14	
finally	6	13	11	18	11	1	60	7
in conclusion	5	9	19	39	6	4	82	9
total	46	96	81	163	58	19	463	39

The frequencies are significant between native and Korean L2 writers, as the Korean students overused these transitionals (χ^2=49.0, p=.003), which can lead to writing that sounds overly artificial or formulaic. This effect holds true for all ESL levels compared to the L1 corpus; students at each ESL level overused these markers compared to the L1 writers. This is apparently an artifact of prescriptive English pedagogy in Korea, particularly when learning essay writing for standardized English tests. Korean students taking TOEFL and TOEIC preparation courses in Korea are taught to use these conjunctions to structure their essays clearly and to make sequencing explicit for better essay test scores. This carries over into their academic writing, as they are not taught that such conjunctions are less common in academic writing. Also, while more of them used the forms that tend to be more North American in style (*first, second, third*), some used the more British sounding forms (*firstly, secondly, thirdly*), and some writers vacillated between the North American and British forms within the same essay, suggesting that they are unaware of this stylistic difference between American and British. Some also use *first of all,* not realizing that this is colloquial style.

In addition, several pragmatically awkward attempts were made with *next*: *next things* (1

token, IEP level), *next thing* (1 token, level 405), *the next* (1 token, 401), and *the next step* (2 tokens, from one UG writer). Several awkward forms of *in conclusion* appeared, such as *as a conclusion* (1 token, UG; 2 tokens, 401; 1 token, 402), *my conclusion...is that...* (2 tokens, 401), *to the conclusion* (1 token, 405), *in the conclusion* (1 token, 401), *for the conclusion*(1 token, 401), and *in a conclusion* (1 token, 402).

5. Topic Shift Markers

The topic of a given sentence (usually the subject) may continue as the topic of a following sentence, or a writer may shift to a related topic such as an example (often with an exemplifier or enumerative marker) – thus, Continue and Smooth-Shift topic flow patterns, in the Centering Theory framework (Walker, Joshi, & Prince, 1998). Other transitional or structural devices may signal more abrupt topic shifts, or resuming a previously mentioned topic (reshift), or similarly, shifting to a new topic that is related to an earlier topic. Such topic resumptions are sometimes managed with sentence-initial reshift markers such as *anyway, as for, as to, as mentioned, regarding, as regards, as concerns*. However, some of these are more informal (*anyway*), and the others are used occasionally in academic writing, but not as often as in informal contexts.

A number of examples of informal topic reshift markers appeared in the Korean L2 Corpus. While *regarding, as regards, with regard to, as concerns* and such are fairly common in academic and formal writing, transitionals like *as for, as to, as in, anyway, speaking of* are more colloquial and informal, and less common in formal writing. While the Korean L2 writers used these informal reshift markers more often than native English writers in the L1 Corpus, who used them very infrequently. Below are Korean L2 examples of these colloquial transitionals, *as for* and *anyway*:

> When serotonin is not released insufficient quantity, depressive symptoms begin to occur. It is very interesting cause, which can shed light the relationship between the light and the depression. As for me, I usually become depressed whenever I stay home long time. This might be the result of the shortage of the sunlight (401 sample).

> The developing this energy faces mechanical difficulties like bulky and cost. But individual person can participate easy way of our mission, For example, if possible, reduce the number of using cars. Let's use mass transportation system. Anyway this is small number of was for saving our life (401 sample).

Infelicitous variations of *as for* included *as for me, and* the redundant *for example, as for me* above. The use of self-referential expressions and examples is less common in most forms of graduate level and research related writing, so their usage as well as using more

colloquial reshift markers can be particularly infelicitous, as they can detract from a formal academic tone in writing. It should be noted, however, that the assigned writing topics may have had an influence on the use of self-referential examples above.

These patterns are likely influenced by Korean topical nouns, which are marked with a topic marker (similar to the *case of* structure in Korean above) affixed directly to nouns just as case markers (nominative, accusative, etc.), e.g., *che-nun,* "*as for me.*" While Korean uses these topic marked noun phrases for minor shifts, English has no such counterpart, and often these shifts are simply unmarked. English sentences in more formal writing simply begin with unmarked topic noun as the sentence subject, while colloquial English more often uses markers like *as to, as for, speaking of.*

One token of *as in* was used as a reshift marker; no tokens of *as to* were found in either corpus. Items such as *concerning* or *regarding* were only examined in clause-initial position, where they would be used as reshift markers; this included an erroneous form *concerning on.* Some of the Korean writers used these expressions with the subjective and informal *me*, while these expressions were rarely used by the L1 writers. The Korean L2 students rarely used the slightly more formal forms like *regarding, as concerns*, but tended to use the *with regard to* and the more colloquial *as for, anyway* for reshift markers. The data here are too sparse for a meaningful statistical analysis, but a clear pattern is evident in the L1 and L2 word counts in Table 8 (no tokens were found in the IEP corpus).

TABLE 8
Topic Reshift Markers

	UG	400	401	402	405	Korean L2 total	L1
as for	1 [me=1]	1 [me=1]	6 [me=4]	1 [me=0]	0	9 [me=6]	2 [me=0]
as in			1			1	0
concerning			1			0	0
as concerns						0	0
concerning on			1	1		2	0
regarding			3	1		4	0
as regards						0	0
with regard to		2	3	2		7	0
anyway						0	5
speaking of						0	0

Another type of expression that Korean L2 writers had difficulty with were discourse anaphoric phrases like *as mentioned, the aforementioned*, where minor morphosyntactic or stylistic errors were made; or they attempted to use the verbs *discussed, told,* or *stated*

instead. The pragmatically or grammatically awkward forms (*) are compared with more correct uses of these words used as topic transitionals below. Correct tokens of *mentioned* include *as mentioned* (6 tokens) and *as mentioned above* (7). Infelicitous uses of *mentioned* include *as I mentioned (3), **as mentioned earlier* (1), **as mentioned before* (2), **as I mentioned above* (5), **as mentioned earlier* (4), **as mentioned before* (3), **as above mentioned* (1), **above mentione*d (1), **already mentioned* (1), **as I have already mentioned* (1), and **as it was already mentioned* (1). Other expressions with the verbs *told, state,* and *discuss* were also used infelicitously.[4] Table 9 shows counts by level, with none in the IEP.

TABLE 9
Discourse Referential Markers in the L2 Corpus

	UG	400	401	402	405	total
mentioned, correct	1	2	9	1		13
mentioned, infelicitous	6	2	25	2		35
*as I told before			1			1
*as states			1			1
*as stated above			1			1
discussed + noun		1	1			2

Awkward uses of *discussed* plus noun include the following example.

> The recommendation of serving size in carbohydrate group is less and protein group is a little bit more recommended according to former <u>discussed</u> clinical evidences (401 sample).

6. *There is/are*

The sentence initial *there is/are* construction is sometimes called presentational *there,* since it used to introduce or present a new, specific item to the discourse (Birner, 1994). It is used more conservatively in academic writing; more commonly, paragraph breaks or other structural devices are often used for such transitions. The Korean L2 writers used *there is/are* significantly more often than native English writers (χ^2=24.63, p<.001), as shown in Table 10.

---

[4] The verb *mention* was also overused and often awkwardly used as a quotative verb, where instead one would use *note, report, observe, state*, etc., in academic writing: Moreover, Liberman <u>mentioned</u> that the hyoid bone of modern humans is similar in size to the hyoid bone of pigs (401 sample).

TABLE 10
There is/are between Both Corpora

	there is	there are	total
IEP	11	23	34
UG	64	143	207
400	33	55	88
401	131	817	948
402	13	15	28
405	5	9	14
Korean L2 total	257	1062	1319
L1	101	73	174

Their use of *there is/are* is influenced by the Korean existential copular verb *issta,* which is a high frequency verb in Korean, and translates as 'there is/are, to be (existential), to have, to exist.' It is thus more challenging for Korean L2 learners to avoid overusing *there is/are* in English, due to the lack of an exact one-to-one equivalent in English, and because Korean writers seem to not know how to manage these topic transitions in English. Predictably, one should see some improvement at higher L2 levels, but an overall trend toward improvement did not hold (χ^2=7.038, p=.218) from lower to higher levels.

There can occur with unaccusative verbs, that is, intransitive verbs of existence, state, and change of state (e.g., *there exists, there appeared, there arose*), which represents a stylistic option that is often preferred over *there is/are* in formal writing. However, no *there +* unaccusative expressions were found in the Korean L2 corpus, suggesting that students were unfamiliar with this option. Koreans may also introduce rather general topics with *there is/are,* such as this example from the L2 corpus:

> Such attitudes can be found when they make products, when they eat or cook and even when they make decisions. <u>There are</u> many countries in the world where mental aspect of human being or saving face is regarded as more important than efficiency or benefits to be expected to produce (401 sample).

Such cases represent a stylistic and pragmatic weakness, in that it does not present a very new, specific item that is sufficiently informative to the reader. Because of this weakness in fulfilling the Gricean informativeness principle, this kind of transitional will come across as unengaging and lacking in formal academic tone. Again, this is probably due to the influence of the multi-purpose existential and presentational Korean verb *issta,* while the English *there is/are* is

more limited to presentation of new and specific items to the discourse[5].

V. GENERAL DISCUSSION

Some of the results and pedagogical implications are summarized below. This moderately sized L2 corpus yielded fairly robust results, which accord with other studies of Korean and other East Asian writers. These results are thus likely generalizable, particularly for connector types that have been previously studied such as contrastive or concessive markers and enumeratives used by Koreans (Kang, 2004, 2005). Problems with other connectors are reported in this study, which have rarely been studied in the past, especially for East Asian learners of English; these include *there is/are, especially*, topic markers, and various infelicitous forms. The corpus data also yielded some information about trends across proficiency levels, which have rarely been studied. The essay writers were mostly new arrivals to the U.S. at the time, and thus can be fairly representative of Korean EFL and ESL students.

1. Writing patterns

Several noticeable patterns emerged from the analysis. First, minor but persistent errors were found with the contrastive markers *but* and *yet*. This likely reflects the lack of direct correspondence between English *but, however, yet,* and various possible Korean equivalents. This also reflects the lack of attention to connectors in English writing instruction in Korea, as many students have indicated to this author that usage of these and other connectors is typically neglected in Korean EFL classrooms. These English markers correspond well to analogues in related languages, so they seem less problematic for ESL/EFL learners from L1 backgrounds such as German, Spanish or French, but the lack of pragmatic correspondence between English and Korean connectors poses more problems for Koreans learning English.

L1 influences affect Koreans' English writing in various other ways, particularly in topic flow and topic management markers. These expressions have received little attention in language pedagogy. Errors with such items are related to pragmatic differences between Korean and English in how they manage information flow and topic shifts. Topic shift markers function

---

[5] Sentence inversion was not examined here, i.e., placement of a prepositional, adverbial, or adjectival phrase sentence initially and immediately before the verb, displacing the subject after the verb (e.g., *From the harbor arose a great stench*, preposed locative; *Most common are locative phrases*, preposed adjective phrase). Such structures are used to perform a shift to a new topic that is closely related to the immediately preceding context (Birner, 1994). These structures were very rare or non-existent in this L2 corpus. Inversion as a transitional device and *there* with unaccusatives are structures would need to be explicitly taught to second language writers. A further difficulty with unaccusatives is that Korean writers incorrectly passivize these verbs (e.g., *was existed, was appeared*); this problem among Koreans (Cowan et al., 2003) also holds true for Japanese writers (Kuno & Takami, 2003; Yusa, 2002).

differently between the two languages, as Korean makes more use of topic-marked noun phrases, while formal English either does not explicitly express such types of shift, or occasionally other uses expressions (e.g., *regarding, with regard to*) for explicit transitions. L1 influence can be seen in the overuse of *there is* and *there are*, and the lack of use of *there* plus unaccusative verbs instead, or inverted sentence structures. ESL/EFL teachers of writing also need to wean students off of overuse of transitional devices like *there is/are* and sentence-initial *especially*, and instead encourage use of exemplifiers like *in particular, particularly*, and the formal Latin abbreviations *i.e.* and *e.g.* For intermediate to advanced writing students, alternatives can be taught such as *there* plus unaccusative verbs (and correct non-passive use of such verbs), paragraph breaks, starting sentences with full noun subjects to introduce new topics, and sentence inversion.

Considerable overuse of the *especially* as a sentence initial connective reflects an apparent transfer error from Korean connectives into English writing. Exemplifiers in argumentation were also overused and often used awkwardly: *in (the) case of*, also representing an L1 influence, and other phrases like *as evidence(s)*, again reflecting L1 influences. Enumeratives (*first, second, third*) were also overused, reflecting instructional influences of test-driven English writing instruction in Korea, rather than L1 influences.

One issue not examined here is the discourse referential *it* versus *that*, when referring to a preceding proposition or concept (e.g., *The researchers failed to control for patients' ages; *it / this led to experimental errors*). In such sentences, *this* refers to a whole preceding clause and proposition, and could also be expressed with a noun paraphrase (*this situation, this error, this problem,* etc.). Korean writers tend to use *it* infelicitously instead of the more appropriate *that*. However, this is more of an anaphoric and cohesive device rather than a connective or transitional, beyond the scope of this study, so it was not examined here. This anaphoric device has been discussed in other texts on teaching ESL/EFL academic writing (Hinkel, 2002, 2004; Swales & Feak, 2004).

2. Pedagogical implications

These problems with contrastive markers and other connector problems above tended not to improve at higher proficiency levels. Korean students are thus at a disadvantage in academic English writing, not only due to significant cross-linguistic differences, but also due to a lack of previous systematic instruction in academic writing style and structure, even at higher levels. Discourse connectives and transitional devices are an area where Korean and other L2 learners could benefit from attention to specific forms and structures. In addition to explicit explanations of English connectors and transitional devices, teachers can use inductive activities, in which students are given examples of stylistically poor texts, or multiple versions of a text using different connectors, so they can notice and infer differences in usage or the stylistic awkwardness of some expressions. Peer editing and revision activities can likewise be used to

build awareness of style and usage. Teachers can draw from corpus samples such as the one used here, or collections of L1 essays and anonymous L2 essays for examples for learning activities. To facilitate inductive learning, teachers can mark up an L2 essay with comments, or highlight problematic expressions and have students deduce the problems to build awareness.

Teaching such expressions can begin at earlier levels of instruction, where students learn paragraph structure and different paragraph forms, e.g., contrast and comparison paragraphs, example paragraphs, and process paragraphs, in the traditional genre based approach. In addition to familiar sentence scrambling and arranging exercises to teach the function of connectors, teachers can highlight the different contrastive markers in contrast and comparison paragraphs. This can be done inductively, e.g., by presenting multiple versions of a paragraph with different contrastive markers in the different versions, to draw students' attention to their differences. The default contrastive marker *but* is a general purpose concessive that places equal weight on both clauses. Students can be presented a sample essay with overuse of *but*, and asked to find the stylistic problems and suggestions for improvement, e.g., by using more variety of conjunctions. Subordinating conjunctions (*though, although*) background the content of their clauses (Hinkel, 2002, 2004), so these markers place secondary importance on the subordinate clause contents and primary focus on the main clause contents. Teachers can also teach students to avoid fragmented or incomplete sentences of subordinate clauses stranded without main clauses. The marker *however* can be taught for more variation and formal style, for slightly stronger contrasts than *but,* particularly as an alternative to the colloquial *but* with a comma. At higher levels, teachers can teach students about using more variety in their various types of connectors through revision exercises, and more contrastive markers such as *in contrast* and *to the contrary,* for stronger contrasts, and the concessive markers *while, yet,* and others. Some markers convey more specific types of contrast; e.g., the informal contrastive *still* and the more formal *nevertheless*, which imply "that's a valid assumption; however, the case is..." (Blakemore, 2000). For upper levels, students can also learn to sometimes use more periphrastic expressions instead of conjunctions to avoid overuse, e.g., *in spite of this, a different view / hypothesis proposed by X, in contrast to this view, in opposition to this* in place of simple contrastive conjunctions (see Crewe, 1990).

For process paragraphs, students can first learn to use enumeratives as a learning scaffold at lower levels, and then process paragraphs at later stages without overly depending on enumeratives. Advanced students also need to learn more formal style in areas like those discussed above: exemplifiers (including *e.g.* and *i.e.*), alternatives to sentence-initial *especially* (e.g., *particularly, in particular*, omitting *especially*, or moving *especially* inside the sentence to modify only a single verb or adjective), avoiding overuse of topic reshift markers or colloquial topic shift markers, and avoiding awkward transfer errors such as **as evidence.* Alternatives to *there is/are* need to be addressed at intermediate and higher levels, i.e., sentences with full noun phrase subjects to introduce new topics, paragraph breaks, sentence inversion, and *there* with

unaccusative verbs, where students also need instruction in not incorrectly passivizing unaccusatives.

Corpus research can inform teachers about problems that need to be addressed, especially in training new teachers or teachers lacking experience with a particular L1 group. Also, teachers' intuitions about problems of particular language learners may not always be reliable (Biber & Conrad, 2001) - they might not notice some usage patterns, or may not understand why certain writing patterns occur. The analytical results, and writing samples from corpora, can also inform materials design. Teachers and materials designers can draw examples from corpora to present students examples of correct and infelicitous usage, for inductive or discovery based learning, error correction, or explicit instruction in writing. They can provide direction for instruction in specific writing expressions, as well as appropriate register - in that academic writing has more nominalizations, derivational nouns, and conjunctive adverbials, and fewer deletions of *that* in complement clauses (Biber et al., 1999; Conrad, 1999). The selection of materials and topics in writing and grammar courses can be informed by corpus based studies. Some advocate providing students output from corpus software for inductive or data driven exercises (Johns, 1991; Johns, 2000), though so much data might be overwhelming for many students, especially at lower and intermediate levels.

VI. CONCLUSION

These data reveal a number of issues that writing teachers need to address in English classrooms. Korean and English are very different in morphosyntax, information structure, and pragmatics, so items that seem synonymous across both languages are functionally quite different. To Korean learners of English (or English learners of Korean), expressions that are seemingly similar in the L2 can have confusingly different functions, such as the contrastive markers, exemplifiers, and focus markers, or can have different pragmatic restrictions on their use in different genres, such as *there* expressions in academic English writing. The issue is compounded by a lack of adequate teaching in academic English writing, including traditional prescriptive and test-driven teaching without learning about the style and pragmatics of different forms of English writing.

Corpus studies like this can be helpful for elucidating the problems that L2 learners face, especially when writing teachers who do not know the students' first language may not know what problems to deal with, or why students manifest such stylistic problems. An L2 corpus can not only provide teachers meaningful information about what issues need to be addressed, but also examples to work on in the classroom. Teachers can use authentic materials for writing instruction, such as L1 essays and L2 writing samples for developing awareness of academic English style and expression. In the areas discussed above, ESL/EFL teachers can point out to

students the stylistic and pragmatic conventions of English academic writing, and the differences that tend to exist between colloquial and academic English, and between the L1 and L2 writing style.

REFERENCES

Agresti, A. (1996). *An introduction to applied categorical data analysis*. New York: John Wiley & Sons.

Aijmer, K. (2002). Modality in advanced Swedish learners' written interlanguage. In S., Granger, J., Hung, & S., Petch-Tyson (Eds.), Computer Learner Corpora, Second Language Acquisition and Foreign Language Teaching (pp. 55–76). Philadelphia: John Benjamins.

Baayen, R. H. & Lieber, R. (1997). Word frequency distributions and lexical semantics. *Computers and the Humanities, 30,* 280-290.

Biber, D. & Conrad, S. (2001). Quantitative corpus-based research: Much more than bean counting. *TESOL Quarterly, 35,* 331-336.

Biber, D., Johansson, S., Conrad, S., & Finegan, E. (1999). *Longman Grammar of Spoken and Written English.* Harlow: Longman.

Birner, B. J. (1994). Information status and word order: An analysis of English inversion. *Language, 70,* 233-259.

Brinton, L. J. (1996). *Pragmatic markers in English: Grammaticalization and discourse functions*. Den Haag, Netherlands: De Gruyter Mouton.

Blakemore, D. (2000). Indicators and procedures: *Nevertheless* and *but*. *Journal of Linguistics, 36,* 463-486.

Conrad, S. M. (1999). The importance of corpus-based research for language teachers. *System, 27,* 1-18.

Cowan, R., Choi, H. E., & Kim, D. H. (2003). Four questions for error diagnosis and correction in CALL. *CALICO Journal, 23*(3), 1-14.

Crewe, W. J. (1990). The illogic of logical connectors. *ELT Journal, 44*, 316-325.

Degand, L., Lefèvre, N., & Bestgen, Y. (1999). The impact of connectives and anaphoric expressions on expository discourse comprehension. *Document Design, 1,* 39-51.

Granger, S. (2003). The international corpus of learner English: A new resource for foreign language learning and teaching and second language acquisition research. *TESOL Quarterly, 37*, 538-546.

Granger, S. (2002). A bird's-eye view of learner corpus research. In S. Granger, J. Hung, & S. Petch-Tyson (Eds.), *Computer learner corpora, second language acquisition and foreign language teaching* (pp. 3-33). Amsterdam/Philadelphia: John Benjamins.

Granger, S. (1999). Use of tenses by advanced EFL learners: Evidence from an error-tagged

computer corpus. In H. Hasselgård & S. Oksefjell (Eds.), *Out of Corpora* (pp. 191-202). Amsterdam: Rodopi.

Granger, S., & Tyson, S. (1996). Connector usage in the English essay writing of native and non-native EFL speakers of English. *World Englishes*, *15*, 17-27.

Hinkel, E. (2004). *Teaching academic ESL writing.* Mahwah, NJ: Lawrence Erlbaum Press.

Hinkel, E. (2002). *Second language writers' text.* Mahwah, NJ: Lawrence Erlbaum Press.

Housen, A. (2002). A corpus-based study of the L2 acquisition of the English verb system. In S. Granger, J. Hung, & S. Petch-Tyson (Eds.), *Computer learner corpora, second language acquisition and foreign language teaching* (pp. 77-116). Amsterdam: John Benjamins Publishing Company.

Kang, J. Y. (2005). Written narratives as an index of L2 competence in Korean EFL learners. *Journal of Second Language Writing, 14,* 259-279.

Kang, J. Y. (2004). Telling a coherent story in a foreign language: Analysis of Korean EFL learners' referential strategies in oral narrative discourse. *Journal of Pragmatics, 36*, 1975-1990.

König, E. (1991). *The meaning of focus particles.* London: Routledge.

Kuno, S., & Takami, K. I. (2003). Remarks on unaccusativity and unergativity in Japanese and Korean. In W. McClure (Ed.), *Japanese/Korean Linguistics, 12*, 280-294. Stanford, CA.: CSLI Publications.

McNamara, D. S., & Kintsch, W. (1996). Learning from text: Effects of prior knowledge and text coherence. *Discourse Processes, 22*, 247-288.

Millis, K. K., Graesser, A. C., & Haberlandt, K. (1993). The impact of connectives on the memory for expository texts. *Applied Cognitive Psychology, 7,* 317-339.

Milton, J. C. P., & Tsang, E. S. C. (1993). A corpus-based study of logical connectors in EFL students' writing: Directions for future research. *Proceedings of a seminar on lexis organized by the Language Centre of the HKUST* (pp. 215-246). Hong Kong: HKUST.

Narita, M., Sato, C., & Sugiura, M. (2004). Connector usage in the English essay writing of Japanese EFL learners. *Fourth International Conference on Language Resources and Evaluation* (pp. 1171-1174).

Ringbom, H. (1999). High frequency verbs in the ICLE corpus. In A. Renouf (Ed.), *Explorations in corpus linguistics* (pp. 191-200). Amsterdam: Rodopi.

Schiffrin, D. (1987). *Discourse Markers.* Cambridge: Cambridge University Press.

Swales, J. M., & Feak, C. (2004). *Academic writing for graduate students*. Ann Arbor, MI: University of Michigan Press

Traxler, M. J., & Gernsbacher, M. A. (1995). Improving coherence in written communication. In M.A. Gernsbacher & T. Givón (Eds.), *Coherence in spontaneous text* (pp. 215-238). Amsterdam: Benjamins.

Yang, W., & Sun, Y. (2012). The use of cohesive devices in argumentative writing by Chinese

EFL learners at different proficiency levels. *Linguistics and Education, 23*, 31-48.

Yoon, J. W., & Yoo, I. W. H. (2011). An error analysis of English conjunctive adjuncts in Korean college students' writing. *English Teaching, 66*(1), 225-244.

Yusa, N. (2002). 'Passive' unaccusatives in L2 acquisition. In Clancy, P. M. (Ed.), *Japanese/Korean linguistics, 11* (pp. 246-259). Stanford, CA: CSLI Publications.

Walker, M., Joshi, A., & Prince, E. (1998). Centering in naturally occurring discourse: An overview. In M. Walker, A. Joshi, & E. Prince (Eds.), *Centering theory in discourse* (pp. 1-30). Oxford: Clarendon Press.

Zhang, M. (2000). Cohesive features in the expository writing of undergraduates in two Chinese universities. *RELC Journal, 31,* 61-95.

Kent Lee

Center for Teaching & Learning

121 Central Plaza

Korean University

Anam-dong 5ga, Seongbuk-gu

Seoul 136-701, Korea

Tel.: 010-5506-2083

E-mail: kentlee@korea.ac.kr, kentlee7@gmail.com