A Corpus-based Analysis of Chinese EFL Learners' Use of Linking Adverbials

Xinfeng Feng* · Mun-Hong Choe**
(Chonnam National University)

Feng, Xinfeng & Choe, Mun-Hong. (2016). A Corpus-based Analysis of Chinese EFL Learners' Use of Linking Adverbials. *The Linguistic Association of Korea Journal* *24*(1), 49-70. This study investigated how the learners of English as a foreign language make use of linking adverbials (LAs) in academic writing, specifically concerned with the developmental aspects of Chinese college students' ability to utilize them in argumentative essays. A corpus-based analysis was conducted with a learner corpus composed of four sub-corpora with respect to each year in college and a control corpus of English L1 writers' essays. The results show that Chinese students use LAs excessively as compared to L1 writers. They overuse causal and sequential LAs across all years of study while gradually approaching the target norms in use of additive and adversative LAs. The frequencies of overused LAs fluctuate over time and then eventually decrease in the fourth year. They produce more various LAs as the amount of exposure and learning experience increases. However, they still tend to rely on a small number of LAs and overuse them in prefabricated patterns.

Key Words: linking adverbials, Chinese, English, corpus analysis, argumentative essays

1. Introduction

Linking adverbials (LAs) are devices for textual cohesion in writing. They are used to denote the relationship between two units of discourse (Biber,

* First author
** Corresponding author

Johansson, Leech, Conrad, & Finegan, 1999). Research shows that there is a strong correlation between felicitous use of LAs and cohesion of writing and when instructed properly, LAs can facilitate students' development in cohesive writing (Chen, 2006; Crewe, 1990; Granger & Tyson, 1996; Liu, 2008; Milton & Tsang, 1993). Recent corpus-based studies have been carried out with learners of varying L1 backgrounds, identifying the problems of overuse, underuse, and misuse by L2 learners. Although there exist a number of studies conducted in EFL settings, for example, with Swedish L1 speakers (Altenberg & Tapper, 1998), Hong Kong students (Bolton, Nelson, & Hung, 2003), Taiwanese speakers (Chen, 2006), and Cantonese speakers (Field & Oi, 1992), few studies have directly investigated Chinese college students. Furthermore, no prior research has examined how the learners' ability to make use of LAs changes over time as their exposure to and experience with the target language increase. Thus, the present study sets out to meet the need for research, viz., Chinese college students' interlanguage characteristics and developmental changes in the use of English LAs.

Since Chinese students write argumentative essays more often than other genres of writing, the current analysis focused on argumentative essays. In addition, timed essays were assumed to be more authentic and reflective of the students' latent competence because they are less likely to be affected by external variables such as model essays in writing reference books. The corpus data for analysis were compiled from undergraduate students in each year of study so that the relationship between their ability to use LAs in a target-like manner and increase in years of study could be observed.

2. Literature Review

Linking adverbials have been referred to by a host of names. According to Liu (2008), researchers use terms such as connective adjuncts (Huddleston & Pullum, 2002), connectives (Finch, 2000), linking adjuncts (Carter & McCarthy, 2006), and logical connectors (Celce-Murcia & Larsen-Freeman, 1999). Moreover, conjunctive adverbials (Bussmann, 1996), conjuncts, connective adverbs, and linking adverbials (Biber, Johansson, Leech, Conrad, & Finegan, 1999) are also

used interchangeably. In this study, we will adopt the term "linking adverbials" because as Liu (2008) mentioned, 'linking' is more comprehensible than 'conjunctive' for general readers and 'adverbial' is more inclusive than 'adverb'.

According to Biber et al. (1999), the primary function of LAs is to make semantic connections between spans of discourse of varying lengths. Celce-Murcia and Larsen-Freeman (1999, p. 519) said that linking adverbials are lexical expressions that may add little or no propositional content by themselves but they serve to specify the relationships among sentences in oral or written discourse, thereby leading the listener/reader to the feeling that the sentences make sense. Understanding and using LAs constitute an important part of communicative competence.

Linguists define and classify linking expressions differently. One of the most widely accepted is Liu's (2008), which is presented below and also the one employed in this study.

Table 1. Taxonomy of Linking Adverbials (Liu, 2008)

Additive	emphatic	*above all, again, also, as well, besides, in addition (to)*
	appositional/ reformulation	*i.e., that is, that is to say, in other words, for example*
	similarity/ comparative	*alternatively, by the same token, correspondingly*
Adversative	proper adversative concessive	*at the same time, however, nevertheless, nonetheless, then, again, though, yet*
	contrastive	*in/by contrast, in fact, in reality, on the other hand*
	correction	*instead, on the contrary, rather*
	dismissal	*admittedly, after all, all the same, anyhow, anyway*
Causal/ Resultative	general causal	*accordingly, as a consequence (of), as a result (of)*
	conditional causal	*thus, all things considered, in such a case/cases*
Sequential	enumerative/ listing	*afterwards, eventually, first(ly), second(ly)*
	simultaneous	*at the same time, in the meantime, meanwhile*
	summative	*all in all, in a word, in short, in sum*
	transitional	*by the way, incidentally*

Over the past two decades, a volume of studies have been conducted on LAs within a comparative research framework. They have been interested in how L2 learners' comprehension and production of LAs differ from L1 controls and what kinds of interlanguage features are observed therein. For example, Altenberg and Tapper (1998) examined Swedish EFL learners' use of LAs in the Swedish component of the International Corpus of Learner English. They found that the overall patterns of LAs used by Swedish students in L2 English writing were comparable to those used by English L1 student writers. However, the students avoided using the resultative and contrastive LAs that were normally used in formal registers. As this might be caused by the students' lack of awareness, they recommended that Swedish EFL students be exposed to a broader range of registers and to a more extensive training in expository writing.

Bolton et al. (2003) compared the use of LAs in Hong Kong EFL students' and native British students' writing to that in a control corpus of published academic papers. They used data extracted from the International Corpus of English (ICE). The results indicate that both Hong Kong EFL students and British students use a small number of LAs in their writing and tend to use them excessively. They offered a list of LAs for the analysis of underuse and overuse in students' writing. In their study, sentence was the basic unit of calculation for the ratio of occurrence per million words, so the frequency per sentence of each tested item was provided. They also argue that the best norm for both native and non-native students' writing is published texts in international English-language academic journals.

Chen (2006) explored advanced EFL learners' use of LAs in a specific field of study using two self-compiled corpora: a learner corpus consisting of 23 final papers contributed by 10 TESOL major students from Taiwan and a control corpus consisting of 10 journal articles from two international TESOL journals. Student writers were found to overuse LAs to a significant degree. In addition, quite a few LAs were inaccurately conceived and implemented by the students. In the same vein, Lei (2012) examined Chinese doctoral students' use of LAs in academic writing. The overall frequency of LAs used by the students was significantly higher than that of professional writers. A total of 33 LAs were found to be overused and 25 to be underused by the students. Also, the

students resorted to a smaller set of LAs than professional writers.

Recently, Li (2014) conducted a corpus-based analysis of LAs used in model essays for students of English and genuine essays composed by professional writers with a self-compiled corpus of model essays and the New York Times Annotated Corpus (NYTAC). The study found that the percentages of the ten most frequently used LAs in the two corpora accounted for 46% and 52% of the total tokens. The model essays used significantly more LAs than L1 writers, especially sequential adverbials.

In a synthetic review of the LA literature, Liu (2008) claimed that most references that touched on the issues at hand lack accuracy and reliability. He underscored the need for more comprehensive research, including frequency, usage, and distribution of LAs with regard to types and registers, in this purview of which the present study is situated. It is further different from the foregoing in that an attempt will be made to investigate not only the frequency and usage of LAs in Chinese students' writing but also the developmental changes in it during four years of study in college. The research questions are stated as follows: (1) How do Chinese college students use LAs in their argumentative essays in comparison with L1 writers? (2) How does their ability to use LAs develop as their learning experience increases during four years of study in college? (3) What are the most frequently used LAs in their writing and how do they differ from L1 writers? (4) Which LAs are overused or underused by Chinese students? The potential causes of the students' non-target-like use of LAs and implications for teaching and learning will be discussed at some length.

3. Methods

A learner corpus (LC) was compiled by extracting Chinese college students' argumentative essays from WECCL 2.0 (Written English Corpus of Chinese Learners 2.0). Since WECCL 2.0 contains data from first-year to fourth-year students in college, it is possible to assess the students' developmental changes through cross-sectional comparisons. A native corpus (NC) was compiled from English L1 speakers' argumentative essays in the Louvain Corpus of Native English Essays. The general characteristics of each corpus are shown in Table 2.

Table 2. The Size of Corpus Data

	NC	LC			
		First	Second	Third	Fourth
Total number of words	149,574	46,483	42,674	47,803	29,787
Total number of essays	176	172	179	174	90
The average essay length	850	270	238	274	331
Total number of LAs	1324	875	867	1025	661

The software Antconc 3.2.4 was used to concordance LAs and to estimate the frequency of occurrence. Since many LAs have more than one meaning and are used in different positions in a sentence, manual work was needed to eliminate unrelated data.

Statistical analyses were conducted to compare the frequencies of LAs in each corpus. Frequency per million words was used for sake of standardized comparison. Bolton et al. (2003) criticized the word-based method to be flawed since it does not take into account the fact that LAs work mostly at the sentential level and beyond. They favored the ratio of number of tokens per number of sentences, instead. However, the sentence-based method has a drawback of its own because, as Chen (2006) pointed out, it puts a particular spin on the results if there is a substantial difference in the average sentence length between the sets of data being compared. As a matter of fact, most previous studies have used the word-based method (Biber et al., 1999; Carter & McCarthy 2006; Granger & Tyson 1996; Liu 2008; Milton & Tsang 1993).

In order to compare the results of the current study with the earlier ones, the word-based method was used. More importantly, the five corpora under investigation exhibited substantial variation in average sentence length and essay length, for example, 270 words per essay in the first-year students' corpus, 331 in the fourth-year corpus, and 850 in NC.

4. Results

4.1. Overall Frequency of Linking Adverbials

Overall, the number of LAs found in LC was far greater than that in NC. In terms of frequency per million words (F/M), Chinese students used LAs more than twice as frequently as did L1 writers. This result is consistent with the findings of Biber et al. (1999), Chen (2006), and Lei (2012) among others.

Table 3. Frequencies of LAs in LC and NC

LAs	LC			NC			χ^2	p
	RF	F/M	%	RF	F/M	%		
Additive	739	4432	22	568	3797	43	7.712	.005*
Adversative	622	3730	18	392	2621	29	30.372	.000*
Causal	1099	6590	32	239	1598	18	466.679	.000*
Sequential	968	5805	28	127	849	10	561.423	.000*
Total	3428	20558	100	1324	8865	100	728.229	.000*

RF: raw frequency; F/M: frequency per million words

The overall difference in the frequency of LAs between LC and NC was statistically significant (χ^2 = 728.229, p = .000). In other words, there was a strong tendency for Chinese students to use a lot of LAs in their writing. Table 3 also presents the raw frequencies, frequencies per million, and percentages of four types of LAs in the two corpora. All the types were used more frequently in LC than in NC (4432 vs. 3797, 3730 vs. 2621, 6590 vs. 1598, and 5805 vs. 849 per million), and Chi-square tests confirmed that the differences were significant.

The largest difference was found in sequential adverbials (5805 vs. 849 F/M). Chinese students used them approximately seven times as frequently as did L1 writers. This is in accord with Li's (2014) observation that LAs are used in model essays much more frequently across all the types and the gap is largest with regard to sequential LAs (3645 vs. 448 per million). It is also noteworthy that the causal type was most frequently used in LC, which was the second least used type in NC. Chinese students used nearly four times more causal LAs than L1 writers. On the other hand, the additive type was used most frequently in NC, but it was used less often than causal and sequential LAs in LC.

4.2. Developmental Changes over Time

In the preceding section, the overall frequency of LAs used by Chinese college students was compared with that of English L1 writers. Another goal of the present study is to see how their command of LAs develops with the increase in learning experience from the first grade to the fourth in college. Table 4 shows the sums of LAs used by Chinese students in each year of study in comparison with L1 writers.

Chinese students across all years produced more LAs than L1 writers in terms of total frequencies (18824/20317/21442/22191 vs. 8865 F/M). There was also a salient tendency for the students to use more LAs in tandem with years of study. A closer look reveals that causal and sequential LAs are largely responsible for the increase of the overall frequency of LAs by time.

Table 4. Frequency Comparisons by LA Type across Years of Study

LAs	LC								NC	
	First		Second		Third		Fourth			
	F/M	%	F/M	%	F/M	%	F/M	%	F/M	%
Add	4647	25	4195	21	4644	21	4096	18	3797	43
Adv	3464	18	3327	16	4372	20	3693	17	2621	29
Caus	6820	36	7264	36	5941	28	6311	28	1598	18
Seq	3894	21	5530	27	6485	31	8091	37	849	10
Total	18824	100	20317	100	21442	100	22191	100	8865	100

Add: Additive; Adv: Adversative; Caus: Causal; Seq: Sequential

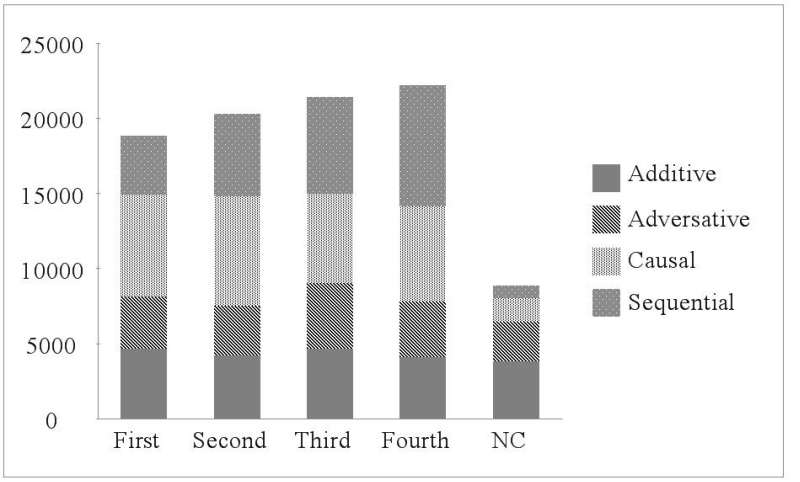
Table 5. Chi-square Tests of Between-groups Differences

Group	Type	First		Second		Third		Fourth	
		χ^2	p	χ^2	p	χ^2	p	χ^2	p
Second	Additive	1.03	.310						
	Adversative	.122	.727						
	Causal	.63	.427						
	Sequential	12.79	.000						
	Total	2.58	.108						
Third	Additive	.00	.995	1.03	.310				
	Adversative	4.97	.026	6.36	.012				
	Causal	2.87	.090	6.05	.014				
	Sequential	30.54	.000	3.42	.064				
	Total	8.18	.004	1.39	.238				
Fourth	Additive	1.25	.264	.41	.839	1.25	.264		
	Adversative	.26	.604	.67	.411	2.06	.150		
	Causal	.713	.398	2.34	.127	.41	.519		
	Sequential	58.11	.000	17.58	.000	6.71	.010		
	Total	10.42	.001	2.98	.084	.48	.487		
NC	Additive	6.42	.011	1.35	.245	6.51	.011	.57	.448
	Adversative	8.95	.003	5.98	.014	36.60	.000	10.23	.001
	Causal	341.94	.000	374.38	.000	258.56	.000	232.37	.000
	Sequential	209.60	.000	386.05	.000	520.85	.000	636.22	.000
	Total	317.98	.000	387.35	.000	488.32	.000	403.83	.000

Table 5 shows the pairwise comparisons of the five corpora under investigation. As for each type of LA, there was no significant difference in the use of additive LAs between second graders and L1 writers (p = .245) and fourth graders and L1 writers (p = .448). First and second graders behaved similarly except for sequential LAs; second graders used significantly more sequential LAs than first graders (3894 < 5530 F/M). Fourth graders in turn used significantly more sequential LAs than second graders. Second graders were different from third graders in respect of adversative and causal LAs; the former used less adversative and more causal LAs than the latter. There was no significant difference between third and fourth graders other than that the latter

used more sequential LAs than the former. These observations imply that Chinese college students' use of LAs undergo significant changes in function of years of study.

Figure 1. Comparison by LA Type across Groups (F/M)



The frequency per million counts and proportions of each type of LA in the corpora are represented in Figure 1. In the L1 corpus, the most frequent type was additive and the least frequent one was sequential. In contrast, those of the learner corpora were sequential and adversative, respectively. Chinese students produced adversative LAs less frequently than the other types regardless of years of study. First and second graders used causal LAs most frequently whereas third and fourth graders used sequential LAs most frequently. Causal and sequential LAs were relatively infrequent in L1 writing.

Chinese students used much more causal and sequential LAs than L1 writers. They overused them through all years of study. Interestingly, with an increase in years of study (and hence the amount of exposure and learning experience), the students used more and more sequential LAs in their writing. Meanwhile, the frequencies of causal LAs alternated throughout. Those of additive and adversative LAs likewise decreased from the first to the second, increased from the second to the third, and then decreased again from the third to the fourth year. On the face of it, there was no apparent evidence for a

continuing improvement in the use of LAs effected by years of study.

Compared to L1 writers, Chinese students overuse causal and sequential LAs. L1 writers tend to use more additive and adversative LAs in their writing. In a point of view, however, the students manifest a considerable change in the use of LAs during the years of study in college. Their use of additive and adversative LAs has eventually become similar to that of L1 writers. A further analysis of individual items demonstrate that the students have made progress with a number of LAs (e.g., *so, then, also, again, besides, of course, what is more, i.e., that is, namely, nevertheless, yet, in reality, on the other hand, instead, despite this/that, as a consequence of, finally* and *in short*). For example, even though *so* is exceedingly overused, its standardized frequencies have gradually declined from the first grade to the fourth. Yet, despite the supposedly increased amount of input and learning through four years of study, their use of sequential LAs has been deviant from the L1 norm to a greater extent.

4.3. Frequently Used Expressions

Table 6 lists the top ten most frequently used LAs in NC and LC with the proportion of each item out of the total number of LAs in the corpus.

Table 6. The Most Frequently Used LAs by Group

NC (%)	LC (%)			
	First	Second	Third	Fourth
also (29)	*so* (33)	*so* (18)	*so* (16)	*so* (9.1)
however (13)	*then* (9.2)	*then* (5.8)	*however* (10)	*second(ly)* (8.2)
yet (4.6)	*also* (8.7)	*however* (5.5)	*second(ly)* (7.7)	*first(ly)* (6.1)
then (4.2)	*however* (8.0)	*second(ly)* (5.2)	*first(ly)* (7.0)	*however* (6.0)
for example (4.0)	*second(ly)* (5.6)	*first(ly)* (4.9)	*also* (6.6)	*third(ly)* (3.5)
therefore (3.6)	*first(ly)* (5.4)	*also* (4.7)	*then* (5.2)	*therefore* (3.4)
so (3.5)	*on the other hand* (4.4)	*in a word* (2.6)	*therefore* (4.6)	*also* (2.7)
thus (3.2)	*besides* (4.2)	*for example* (2.6)	*third(ly)* (3.8)	*for example* (2.7)
though (2.1)	*for example* (4.1)	*third(ly)* (2.4)	*for example* (3.5)	*then* (2.7)
in fact (1.8)	*what's more* (3.9)	*what's more* (2.0)	*thus* (2.9)	*last* (1.7)
Total (69)	Total (86.5)	Total (53.7)	Total (67.3)	Total (43.4)

There are common features shared by both Chinese EFL learners and English L1 speakers. Five items are identical across all the groups: *also, for example, however, so,* and *then*. The ten most frequently used LAs account for over half of the LAs used by 1st-3rd graders and L1 writers. It appears that L1 speakers as well as L2 learners employ a rather small set of LAs in their writing.

The causal conjunctive *so* was found most frequently in Chinese students' writing, although its proportions declined with increase in years of study. L1 writers also used it often (3.5% of the total), but much less than Chinese students. Eight out of the ten were shared by all learner groups: *also, for example, however, so, then, first(ly), second(ly), third(ly)*. The majority were of sequential type. However, L1 writers seldom used sequential LAs. In Lei (2012), the ten most frequent LAs accounted for about 60% of all LAs in the student corpus and the control corpus alike. In Li's (2014) study, among the top ten LAs in model essays, sequential LAs appeared with the highest frequency (6 out of 10). *First(ly), second(ly)* were extremely frequently used in model essays, but rarely in NYTAC.

4.4. Linking Adverbials Overused and Underused in the Learner Corpus

Since there is in practice no precise definition of overuse and underuse, previous studies arbitrarily set the criterion to be over 10 per million words in frequency of occurrence (e.g., Chen, 2006; Lei, 2012). However, Li (2014) argues that the pairs of LAs in question whose difference reaches the level of statistical significance be regarded as deviant to a considerable degree. The current study adopts Li's approach to the assessment of overuse and underuse. Items with less than 10 observed cases were not subject to statistical tests.

As seen in Table 7, first-year students overused a total of 15 LAs. The most overused item was *so*, with a difference of 4312 F/M from the L1 corpus. *Then, second(ly)* and *first(ly)* were the next common LAs with a difference of 944, 763, and 721 F/M respectively from the L1 corpus. Seven of the overused LAs were sequential (*second(ly), first(ly), in a word, last(ly), third(ly), meanwhile,* and *first of all*), 2 were additive (*besides* and *of course*), 3 were adversative (*on the other hand, at the same time,* and *after all*), and 3 were causal (*so, then,* and *as a result*).

Table 7. LAs Overused by First Graders

	First Grader		NC		χ^2	p
	RF	F/M	RF	F/M		
so	215	4625.34	47	314.22	493.867	.000
then	61	1312.30	55	367.71	53.511	.000
second(ly)	37	795.99	5	33.43	92.262	.000
first(ly)	36	774.48	8	53.49	82.159	.000
besides	28	602.37	13	86.91	45.065	.000
on the other hand	29	623.88	19	127.02	35.767	.000
in a word	16	344.21	0	N/A	51.479	.000
last(ly)	16	342.28	4	26.74	35.032	.000
third(ly)	15	322.7	2	13.37	39.136	.000
of course	21	451.78	20	133.71	17.159	.000
at the same time	16	344.21	10	66.86	20.573	.000
as a result (of)	16	344.21	20	133.71	8.559	.003
meanwhile	12	258.16	1	6.69	33.824	.000
after all	11	236.64	8	53.48	12.277	.000
first of all	10	215.13	9	60.17	8.778	.003

Among these LAs, only 5 were used by L1 writers over 100 F/M: *so*, *then*, *on the other hand*, *of course*, and *as a result (of)*, none of which were sequential. Indeed, all the sequential LAs overused by the students were extremely rarely used by L1 writers, less than 10 out of 150,000 words in raw frequency. For example, L1 writers never used the expression, *in a word*.

Second-year students overused a total of 14 LAs. Again, *so* was ranked first, with a bit smaller difference of 4044 F/M from the L1 corpus. As in the case of first-year students, sequential LAs such as *first(ly)*, *second(ly)* and *third(ly)* were among the most overused items. Of the 14, 6 were sequential, 4 were additive, 2 were adversative, and 2 were causal LAs. In comparison with first-year students, 2 sequential items (i.e., *in a word* and *meanwhile*) were not found any more while another sequential item, *to sum up*, appeared for the first time.

Table 8. LAs Overused by Second Graders

	Second Grader		NC		X^2	p
	RF	F/M	RF	F/M		
so	186	4358.42	47	314.22	448.640	.000
second/secondly	67	1570.04	5	33.43	209.394	.000
first/firstly	50	1171.67	8	53.49	137.642	.000
then	59	1382.57	55	367.71	57.702	.000
third/thirdly	50	1171.67	2	13.37	164.732	.000
besides	19	445.24	13	86.91	25.615	.000
what's (is) more	19	445.24	8	53.49	36.286	.000
last/lastly	18	423.43	4	26.74	45.287	.000
moreover	11	257.77	5	33.43	20.079	.000
of course	17	398.37	20	133.71	12.086	.001
on the other hand	14	328.06	19	127.02	7.819	.005
at the same time	10	234.33	10	66.86	10.235	.000
to sum up	10	234.33	0	N/A	35.052	.000
first of all	11	257.77	9	60.17	12.462	.000

Another point of interest is that 2 new additive LAs (i.e., *what's (is) more* and *moreover*) began to occur frequently. Also, 1 adversative (i.e., *after all*) and 1 causal (*as a result of*) LAs were not used to an excessive degree.

Third-year students overused even more LAs (22 in total) than first and second-year students. The largest statistical difference was again found with *so*, but the size became considerably smaller than before. Similarly, formulaic sequential LAs such as *first(ly)*, *second(ly)*, and *third(ly)* were still used excessively. Among the overused LAs, 7 were sequential, 6 were additive, 5 were adversative and 4 were causal. Perhaps the most striking feature of this stage is that the learners try to produce a number of new LAs that they did not routinely use at the earlier stages.

Table 9. LAs Overused by Third Graders

	Third Grader		NC		χ^2	p
	RF	F/M	RF	F/M		
so	135	2824.09	47	314.22	247.709	.000
second/secondly	67	1401.58	5	33.43	185.956	.000
first/firstly	61	1276.06	8	53.49	154.943	.000
however	95	1987.32	175	1169.99	17.715	.000
then	45	941.36	55	367.71	23.541	.000
third/thirdly	33	690.33	2	13.37	93.637	.000
therefore	40	836.78	48	320.91	21.631	.000
what's (is) more	24	502.06	8	53.49	44.968	.000
all in all	20	418.38	1	6.69	57.716	.000
in a word	21	439.30	0	N/A	65.715	.000
on the other hand	23	481.14	19	127.02	21.352	.000
at the same time	17	355.62	10	66.86	22.086	.000
besides	15	313.79	13	86.91	13.146	.000
in addition (to)	18	376.55	15	100.29	16.539	.000
of course	17	355.62	20	133.71	9.518	.002
for example	30	627.58	52	347.65	6.835	.009
for instance	17	355.63	11	73.54	20.322	.000
though	22	460.22	28	187.2	10.662	.001
actually	13	271.95	1	6.68	35.939	.000
first of all	17	355.62	9	60.17	24.009	.000
to begin with	14	292.87	1	6.69	39.043	.000
otherwise	10	209.2	9	60.17	8.358	.004

The LAs come in various types: *however, therefore, all in all, in addition, for example, for instance, though, actually, to begin with,* and *otherwise,* the majority of which are additive and adversative LAs. They are also commonly used by L1 writers. This implies that the students are making progress in the control of additive and adversative LAs while being more conservative with sequential LAs.

Table 10. LAs Overused by Fourth Graders

	Fourth Grader		NC		χ^2	p
	RF	F/M	RF	F/M		
so	80	2685.73	47	314.22	197.441	.000
second/secondly	72	2417.16	5	33.43	328.923	.000
first/firstly	54	1812.87	8	53.49	222.517	.000
third/thirdly	31	1040.73	2	13.37	142.524	.000
therefore	30	1007.15	48	320.91	26.911	.000
last/lastly	15	503.25	4	26.74	53.322	.000
in conclusion	14	470.00	0	N/A	70.306	.000
for example	24	805.72	52	347.65	12.306	.000
then	23	772.15	55	367.71	9.347	.002
besides	13	436.43	13	86.91	20.937	.000
moreover	10	335.71	5	33.43	27.144	.000
all in all	10	335.71	1	6.69	43.852	.000

Fourth-year students overused the least number of LAs (12 in total), whereof 6 were sequential, 3 were additive, 3 were causal, and none were adversative. This indicates that from the perspective of individual LAs, Chinese students' performance has become more comparable to that of L1 writers in function of time. The most overused item was *second(ly)*, with a much smaller difference of 2384 F/M. *So* and *first(ly)* were the second and third most overused LAs. All the items but *in conclusion* were found in the earlier years as well. The expression *all in all* continued to be overused. In fact, summative LAs such as *in a word, to sum up, all in all,* and *in conclusion* were almost never used by L1 writers.

On the whole, Chinese students overused *so, then, besides, firstly, secondly,* and *thirdly* across all years of study in college. The causative *so* was exceptionally overused by Chinese students. Another causative *then* was also highly overused throughout. However, their occurrences gradually decreased from the first year to the fourth. In the meantime, sequential LAs were used even to a greater extent with an increase in years of study. This fact gives a partial account to the increased proportion of sequential LAs along with years of study.

Conversely, Chinese students also underuse certain LAs.

Table 11. LAs Underused by Chinese EFL Students

LAs	First Grade		NC		χ^2	p
	RF	F/M	RF	F/M		
again	0	N/A	10	66.87	31.048	.000
yet	3	64.54	61	407.82	12.806	.000
furthermore	1	21.51	9	60.18	23.380	.000
i.e.	0	N/A	4	26.74	12.418	.000
in reality	0	N/A	6	40.11	18.627	.000
rather	0	N/A	8	53.48	24.837	.000
	Second Grade		NC			
yet	4	93.73	61	407.82	9.691	.002
eventually	0	N/A	17	113.66	4.851	.028
in reality	1	23.43	6	40.11	12.994	.002
again	1	23.43	10	66.87	24.126	.000
i.e.	0	N/A	4	26.74	11.400	.001
rather	0	N/A	8	53.48	22.802	.000
	Third Grade		NC			
yet	3	62.76	61	407.82	13.307	.000
eventually	0	N/A	17	113.66	5.434	.020
i.e.	0	N/A	4	26.74	12.770	.000
in reality	0	N/A	6	40.11	19.156	.000
rather	1	20.92	8	53.48	20.962	.000
	Fourth Grade		NC			
yet	4	134.29	61	407.82	5.131	.024
rather	0	N/A	8	53.48	15.917	.000

First-year students underused 6 LAs, among which 3 were additive (*again*, *furthermore*, and *i.e.*) and 3 were adversative (*yet*, *rather*, and *in reality*). Second-year students underused 6 LAs too. Two of them were additive (*again* and *i.e.*), 3 were adversative (*yet*, *rather*, and *in reality*), and 1 was sequential (*eventually*). Third-year students underused 5 LAs. One was additive (*i.e.*), 3 were adversative (*yet*, *rather*, and *in reality*), and 1 was sequential (*eventually*). Fourth-year students underused 2 LAs, both of which were adversative (*yet* and *rather*).

The most underused type was adversative. This echoes Li's (2014) finding that adversative LAs are not readily used by L2 learners. Lei (2012) also reported that in the academic writing of Chinese doctoral students, adversative adverbials were most problematic and nearly half of the LAs underused by the students were adversative. According to Biber et al. (1999), adversative LAs import a more complex relationship between discourse units, which is difficult for unskilled student writers to manipulate.

5. Discussion

Most notably, Chinese students overuse casual and sequential LAs in their writing and they use more and more sequential LAs as they spend more years in college. They rely heavily on a small repertoire of LAs in the first year (87% of the total instances are occupied by 10) and become able to use more various LAs as time goes on. The ten most frequently used LAs account for 43% of the total tokens in the fourth year. This finding is at odds with Liu (2008) in which several frequently used LAs accounted for well over 50% of the total LAs used in the student corpus.

Although second graders overuse additive and adversative LAs in their writing, the frequencies per million of the overused items are lower than those of first graders and thus are closer to the L1 norms. It can therefore be said that second graders are better able to use additive and adversative LAs than first graders. Moreover, there are improved aspects on the use of causal adverbials with increase in years of study, especially from the second grade to the third. It is evident that third graders are better able to use additive and causal LAs than first graders. Fourth graders' use of additive LAs is indeed more target-like than the other groups. Developmental changes on the use of additive and adversative LAs along with increase in years of study are not so obvious. However, the students' ability to use adversative LAs seems to develop at a later stage, particularly from the third grade to the fourth. Besides, third and fourth-year students are more competent at using causal adverbials than first and second-year students.

Developmental changes are also observed as to individual LAs. For example, fourth graders do not excessively use the expression *that is* any more unlike the

other student groups. Although they invariably overuse LAs as compared to L1 writers, the frequencies per million of the redundant LAs go into gradual decline as a function of time. These LAs are *again, as well, in addition to, furthermore, of course, what's more, i.e., that is, in other words, for instance, nevertheless, though, actually, on the other hand, instead, as a consequence (of), naturally, then, to begin with, finally,* and *meanwhile.* To illustrate, the frequencies per million of the adversative *nevertheless* and the causal *so* decline steadily from the first grade to the fourth and eventually become comparable to the L1 norms.

Years of study in college also allow the students to start using new LAs that they did not regularly use before. These LAs include *again, further, not to mention, i.e., in reality, rather, afterwards and in sum/summary.* For example, the usage frequencies of the additive *again* from the first grade to the fourth are 0, 23, 84, and 67, the ultimate figure being identical with that of L1 writers. The sequential LA *in sum/summary* first appears in the third year and is extensively used in the fourth year, the production rate of which is similar to that of L1 writers. Moreover, as their years of study increase, they begin using *furthermore* from the second grade, *again* from the third grade and *i.e.* and *in reality* in the fourth grade to the extent that L1 writers do. Chinese students overuse a number of LAs, of which about half are of sequential type. At the same time, they also underuse some LAs, and most of them are of adversative type.

One possible reason for Chinese students' overuse of LAs is that they try to achieve surface logicality in their writing by means of adding explicit LAs. In Liu (2008), advanced-level students overused LAs in academic writing and an in-depth analysis revealed that the sentences conjoined by the superfluous LAs were actually loosely connected and lacked coherence. Another possibility is that Chinese students are predisposed to mechanically imitate model essays in English writing textbooks. These textbooks do not use LAs with discretion and so misguide the students to blindly accept the contrived usage of LAs. In Li's (2014) study, model writings presented in textbooks use far more LAs than are expected to occur in L1 writing, especially sequential ones. Chinese students usually learn English as a foreign language in order to pass high-stakes tests such as TEM-4 (Test for English Majors Band 4) and TEM-8 (Test for English Majors Band 8). To gain a better score in English writing, they try to imitate the set patterns of model essays. The use of LAs in such reference materials does

not accurately represent authentic language. Consider the following example essay from the third-year student corpus:

> Not until recently, we human beings have spared no effort to speed up our economy development, which seems to keep pace with the whole world's advancement. **However**, this have brought great pains to our environment as well. **Subsequently**, this post a big threat to our health, since a variety of disease break out. Here it comes we need do something to get rid of such a "throw-away society" and better our living condition. **Firstly**, people especially the customers ought to get rid of their trouble-fearing conception. The one-off products, such as chopsticks and canteens, are popular with more and more people now, which in fact bring us harm but no good, for a large amount of chemicals are their ingredients. The merchandise should make sure their dishes are clean and no one-off materials are used. **Secondly**, while shopping, do not use plastic bags, whose ingredients are difficult to be degraded. Paper bags are better choice, they do less harm to our environment, and they have the same function. **In a word**, don't care so much of the economical interests, we'd better follow the sustainable development, and we're obliged to better our living condition.

In this essay, the student argues in a layered discourse structure using a set of formulaic LAs such as *firstly*, *secondly*, and *in a word*. Of course, there are other writing samples with more complex structures and linking expressions, but students generally prefer to use simpler ones, resulting in overusing of sequential LAs and underusing of more intricate ways of building coherence. Conventional teaching practice in the classroom may also mislead Chinese students to use LAs excessively in their writing. According to Crewe (1990), many teachers in China encourage their students to write a composition on the basis of linking devices rather than information units interconnected with each other. Moreover, most textbooks offer lists of non-equivalent linking expressions as if they were exact alternatives.

Therefore, instructors and materials should provide a developmental guidance for the function and judicious use of LAs with exemplary input in a broad range of genres, registers, and discourse settings (cf. Crewe, 1990; Granger & Tyson, 1996). Since Chinese students tend to underuse adversative LAs, emphasis should also be put on those underused items as well. Finally, it goes without saying that the practice of test-oriented teaching and learning needs to be improved.

References

Altenberg, B., & Tapper, M. (1998). The use of adverbial connectors in advanced Swedish learners' written English. In S. Granger (Ed.), *Learner English on computer* (pp. 80 – 93). London: Addison-Wesley, Longman.

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

Bolton, K., Nelson, G., & Hung, J. (2003). A corpus-based study of connectors in student writing: Research from the International Corpus of English in Hong Kong (ICE-HK). *International Journal of Corpus Linguistics, 7*(2), 165-182.

Bussmann, H. (1996). *Routledge dictionary of language and linguistics*. London: Routledge.

Carter, R., & McCarthy, M. (2006). *Cambridge grammar of English*. Cambridge, UK: Cambridge University Press.

Celce-Murcia, M., & Larsen-Freeman, D. (1999). *The grammar book*. Boston: Heinle & Heinle.

Chen, C. W. Y. (2006). The use of conjunctive adverbials in the academic papers of advanced Taiwanese EFL learners. *International Journal of Corpus Linguistics, 11*(1), 113-130.

Crewe, W. J. (1990). The illogic of logical connectives. *ELT Journal, 44*(4), 316–325.

Field, Y., & Oi, Y. L. M. (1992). A comparison of internal conjunctive cohesion in the English essay writing of Cantonese speakers and native speakers of English. *RELC Journal, 23*(1), 15-28.

Finch, G. (2000). *Linguistic terms and concepts*. NY: Macmillan Press.

Granger, S., & Tyson, S. (1996). Connector usage in the English essay writing of native and non-native EFL speakers of English. *World Englishes, 15*(1), 17-27.

Huddleston, R., & Pullum, G. K. (2002). *The Cambridge grammar of English language*. Cambridge, UK: Cambridge University Press.

Lei, L. (2012). Linking adverbials in academic writing on applied linguistics by Chinese doctoral students. *Journal of English for Academic Purposes, 11*(3), 267-275.

Li, A. J. (2014). Do model essays in English writing textbooks potentially misguide EFL learners' use of linking adverbials? A corpus-based study. *The Journal of Studies in Language, 29*(4), 907-926.

Liu, D. (2008). Linking adverbials: An across-register corpus study and its implications. *International Journal of Corpus Linguistics*, *13*(4), 491-518.

Milton, J. C., & 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* (pp. 215-246). Language Centre, HKUST, Hong Kong.

Xinfeng Feng
Department of English Education, Chonnam National University
Yongbong-ro 77, Buk-gu
Gwangju, 61186, Korea
Email: xinfeng0720@gmail.com

Mun-Hong Choe
Department of English Education, Chonnam National University
Yongbong-ro 77, Buk-gu
Gwangju, 61186, Korea
Phone: 062) 530-2435
Email: munhong@jnu.ac.kr

Received on January 2, 2016
Revised version received on March 18, 2016
Accepted on March 30, 2016