THEMES IN SCIENCE AND TECHNOLOGY EDUCATION
Volume 3, Number 1-2, Pages 139-157

A corpus-based study on Asian learners' use of English linking adverbials

Shin'ichiro Ishikawa
iskwshin@gmail.com
Kobe University, Japan

Abstract

ICT influences various phases of language studies and education. Its application for applied linguistics has been mainly attempted in the field of corpus linguistics, which describes how native speakers (NS) use the language and how its use by non-native speakers (NNS) deviates from the NS norm. In the current study, we focused on the English linking adverbials (LA), which plays an important role in the logical cohesion of a text. By analyzing a newly-compiled international learner corpus to compare LA uses by English NS and Asian NNS, including Japanese learners of English (JLE) and their Chinese counterparts (CLE), we obtained several noteworthy findings. First, the gap between NS and Asian NNS is rather limited in terms of the quantity of LA used in essay writing. Second, Asian NNS tend to overuse addictive types of LA. Third, Asian NNS overuse LA items concerning the introduction of additional information and intensification of meaning, while underusing LA items concerning the introduction of parallel information and sequential introduction of information. Finally, JLE, CLE, and NS use major LA items in their own ways, which are hardly influenced by L2 proficiency.

Introduction

ICT and applied linguistics

Although linguistic studies have traditionally resorted to native speakers' intuition, which was believed to show every aspect of a language, "intuition-based grammar" and "observation-based grammar" are not necessarily the same (Aarts, 1991). Closer observation of language use reveals that *what we can say* and *what we actually say* differ substantially. The former could be studied by intuition alone, but the latter needs to be examined via real language data.

The development of ICT has enabled the compilation of large-scale corpora, which are defined as authentic sets of electronically-stored language data, and their

quantitative analysis. By using a concordancer or a specific computer program for linguistic analysis, we can easily determine the frequency of particular words and expressions. As Svartvik (1990) mentions, computer/corpus-based applied linguistics yields "more academic cross-fertilization and fresh approaches to old problems" and leads to "a better understanding of the complexities of natural language and the marvel of human language processing."

There are two major directions in corpus linguistics. One is a study based on corpora of native speakers (NS), which targets a detailed linguistic description of a target language; and the other is a study based on the so-called learner corpora, collecting language use by non-native speakers (NNS), which aims to elucidate the NS/NNS gap and contribute to better language teaching. More and more researchers in the fields of TEFL (Teaching English as a Foreign Language) and SLA (Second Language Acquisition) have gradually recognized the potential pedagogic value of learner corpora.

Linking Adverbials

Although there are varied grammatical items where NS/NNS gaps are observed, we will focus on English linking adverbials. Linking adverbials (LA) explicitly indicate the semantic relationship between textual segments and play a crucial role in making a text logically cohesive. Therefore, it is vital for learners of English, whose writing or speech is often said to be lacking in logical lucidity, to use LA qualitatively and quantitatively in an appropriate way. In the current study, we will compare the frequencies and usages of LA adopted by English NS and two kinds of Asian NNS, namely, Japanese learners of English (JLE) and Chinese learners of English (CLE).
LA vary significantly in length, form, and function. According to Sinclair (2005), LA can be structurally divided into two types: lexical LA (e.g. *therefore*) and phrasal LA (e.g. *as a result*). LA are also functionally sub-categorized into several groups. Quirk and Greenbaum (1973) classifies LA, which they call "conjuncts," into twelve functional groups. Biber et al. (1999) classifies them into six groups. More recently, Sinclair (2005) and Carter & McCarthy (2006) subdivide LA, which they call *linking adjuncts*, into seven and nine groups respectively (Table 1).

Based on detailed analysis of large-scale NS corpora, Biber et al. (1999) yields several interesting facts concerning LA usage: (i) the genre where LA occurs most often is academic texts, followed by conversation, fiction, and news; (ii) the most common functional type of LA is results and inference, both in academic texts and conversations, followed by apposition in the former, but contrasts and concession in the latter; (iii) in academic texts, lexical LA are used quite often, but prepositional linking phrases are also used to a certain extent, while in conversation almost all LA take the form of the former; (iv) the top three LA in academic texts are *however*, *thus*

and *for example*, but in conversations, *so*, *then*, and *though*; and (v) in academic texts LA tend to appear in the initial and middle positions, as opposed to the initial and final positions in conversations.

Among the categories summarized in Table 1, we are going to adopt the framework proposed by Carter & McCarthy (2006) where LA are subcategorized into nine functional types: additive, concessive, contrastive, inference, listing, meta-textual, resultative, summative, and time.

Table 1. Sub-categorizations of LA

Quirk & Greenbaum (1973)		**Biber et al. (1999)**	
Functions	**Items (e.g.)**	**Functions**	**Items (e.g.)**
antithetic	instead	apposition	namely
apposition	namely, e.g.	contrast/ concession	though, alternatively
concessive	nevertheless	enumeration and addition	first, additionally
enumerative	first, second	result/inference	therefore, thus
inferential	else, otherwise	summation	altogether, overall
equative	equally, likewise	transition	incidentally
reformulatory	rather		
reinforcing	also, furthermore		
result	consequently		
summative	then, in conclusion		
temporal transition	meantime, meanwhile		
transitional	by the way		
Sinclair (2005)		**Carter & McCarthy (2006)**	
Functions	**Items (e.g.)**	**Functions**	**Items(e.g.)**
addition	also, at the same time	additive	also, likewise
causes	consequently, so	concessive	anyway, though
conjunctions	and, but, yet	contrastive	rather,
contrasts and alternatives	all the same, even so	inference	then, in that case
ordering points	first, second	listing	firstly, lastly
parallel	again, equally	meta-textual	namely, so to speak
sequence in time	afterwards, first	resultative	so, therefore
		summative	overall, in short
		time	eventually, then

Literature Review

Many studies aim to illuminate the general linguistic differences seen in writing by NS and NNS. Most previous studies in the field are based on a comparison of the International Corpus of Learner English (ICLE) and the Louvain Corpus of Native English Essays (LOCNESS), both of which were compiled on the initiative of Professor Sylviane Granger of Louvain University. To date, topics such as frequent vocabulary (Ringbom, 1988), intensification of adjectives (Lorenz, 1988), pharicons (i.e. phraseological units) (De Cock et al., 1988), direct questions (Vitanen, 1988), writer-reader relationships (Petch-Tyson, 1988), complementary clauses (Biber & Reppen, 1998), modality (Aijmer, 2002), verb system (Housen, 2002), and small words (Hasselgren, 2002) have been analytically discussed.

Also, several studies have focused particularly on L2 learners' use of LA. Altenberg & Tapper (1998), which compared the usages of conjuncts by NS and Swedish NNS at an advanced level, discovered that (i) Swedish NNS generally use a greater variety of conjuncts but the overall frequency is relatively lower; (ii) they overuse appositive type of conjuncts (e.g.: *for example*) while underusing resultive (e.g.: *as a result*) and contrastive (e.g.: *rather*) types; (iii) they overuse conjunct items such as *furthermore*, *for instance*, *still*, and *of course* while underusing those such as *hence*, *therefore*, *thus*, *however*, *though*, and *yet*; and (iv) they tend to use conjuncts mainly in the medial position, while NS favor the clause-initial position. Comparing writing by Norwegian advanced learners and English NS, Eia (2006) revealed that Norwegian NNS use fewer tokens and more types of LA, which supports the findings of Altenberg & Tapper (1998). Narita, Sato, & Sugiura (2003) compared the use of LA by NS and JLE, concluding that Japanese NNS tend to overuse the additive, enumerative, appositive, and stance types of adverbials, while underusing the concessive type.

Although the findings of the previous studies are linguistically and pedagogically noteworthy, most of the corpora used for them were inappropriately controlled in terms of writing conditions. For example, the total number of topics in ICLE amounts to as many as 922, which may have significantly influenced the lexical distribution of the corpus (Ishikawa, 2008). In the present study, therefore, the author uses a new controlled learner corpus to reconsider how the gap between NS and Asian NNS influences the use of LA in their argumentative essays.

Research Design

Aim and Research Questions

The overall aim of the present study is to answer the following research questions: (1) Do Asian NNS use fewer but more varied LA compared to NS? (2) Which functional type of LA is overused or underused by Asian NNS? (3) Which LA items are

overused or underused by Asian NNS? And (4) To what extent are JLE at varied L2 proficiency levels, CLE, and NS close to each other in terms of LA usage? By adopting a research approach called contrastive interlanguage analysis (CIA) (Granger, 1998; Granger, 2002), we aim to quantitatively elucidate how each of the writer groups differs in terms of LA usage.

Data

For the purpose of comparative analysis, we used the Corpus of English Essays Written by Asian University Students (CEEAUS) (the 200904 version), which the author has recently released. CEEAUS collects English argumentative essays written by NS (37,173 tokens), JLE (169,654 tokens), and CLE (20,367 tokens). As mentioned above, the writing conditions of most existing learner corpora are not strictly controlled, meaning it is not always clear whether the observed difference is really due to the difference in writer-groups or the variation in topics of the essays.

Unlike other major learner corpora, for CEEAUS, writing conditions such as topic, length, time, and dictionary use are strictly controlled. In particular, the number of topics is limited to only two: one being "It is important for college students to have a part time job" and the other "Smoking should be completely banned at all the restaurants in the country." In both cases, writers were required to show clearly whether they agreed or disagreed with the proposed statements and also support their claims with appropriate examples. Limiting the number of topics makes the content of the corpus much more lexically homogeneous, which enables us to conduct a robust comparison among different writer-groups (Ishikawa, 2008).

CEEAUS also holds detailed L2 proficiency data of JLE, who are classified into four levels: Lower (-495 in TOEIC® test), Middle (500+), Semi-Upper (600+), and Upper (700+). This proficiency-based subdivision makes it possible to compare NNS at different L2 proficiency levels as well as NNS with different L1 backgrounds. The sizes of the sub-modules in CEEAUS are shown in Table 2.

Table 2. Tokens of the sub-modules of CEEAUS

Module	NS	JLE	CLE	JLE_L	JLE_M	JLE_S	JLE_U
Tokens	37,173	169,654	20,367	17,580	85,614	57,452	9,008
Types	3,797	4,800	1,818	1,642	3,608	2,845	1,206

Scope

The total number of LA items introduced in Carter & McCarthy (2006) amounts to one-hundred nineteen, but if removing those in multiple categories (e.g.: *of course* belonging to resultive and concessive; *otherwise* to contrastive and inference) and those whose forms cannot be lexically embodied (e.g. *a, b, c*), the total decreases to ninety eight.

We examined the frequency of ninety eight LA items used by three different writer-groups: NS, JLE, and CLE. Also, for mutual comparison, we converted the raw frequency to the adjusted frequency per 10,000 tokens (See the appendix). Finally, based on the average of the three adjusted frequencies, the top twenty items were chosen for detailed analysis.

Table 3. Top twenty LA items used for detailed analysis

Rank	LA items	Functions	Av. of Freq. (Adjusted)
1	also	Addictive	36.55
2	only	Concessive	28.85
3	however	concessive	11.92
4	too	addictive	11.08
5	now	meta-textual	10.75
6	well	meta-textual	7.77
7	still	concessive	6.02
8	though	concessive	5.27
9	moreover	addictive	3.92
10	finally	listing	3.62
11	next	listing	2.87
12	firstly	listing	2.45
13	rather	contrastive	2.31
14	instead	contrastive	2.08
15	yet	concessive	1.79
16	furthermore	addictive	1.73
17	indeed	meta-textual	1.30
18	eventually	time	0.96
19	anyway	concessive	0.82
20	lastly	listing	0.69

Methodology

First, we counted up the respective totals for the types and tokens of all LA items used by NS and Asian NNS groups. The number of types represents the variety of LA, while that of tokens represents the quantity of LA. When comparing tokens, we focused on the adjusted frequencies (RQ1).

Subsequently, we comparatively examined the cumulative frequencies of all LA items belonging to each of the functional types (RQ2). Next, focusing on the top twenty LA items, we compared their frequencies in order to identify items overused or underused by JLE and CLE (RQ3). When comparing the frequencies of the three writer-groups, we conducted a chi-square test with Bonferroni correction. If the *p*-value was smaller than .016 (i.e. .05 / 3), the difference was regarded as statistically significant.

Finally, we conducted correspondence analysis, which is an explanatory statistical technique developed to analyze multi-way tables containing correspondence between rows (item 1) and columns (item 2), to classify the writer-groups and individual LA items (RQ4). Seven kinds of writer-groups (NS, JLE, CLE; JLE_Lower, JLE_Middle, JLE_Semi-upper, JLE_upper) are regarded as item 1 and the top twenty LA items as items 2. By applying correspondence analysis, we examined the affinity among different writer groups in terms of LA use and the relationship between each of the writer groups and individual LA items.

Results and Discussion

Quantity and variety

The total numbers of tokens for all LA items used by NS, JLE, and CLE are 517 (139.1 per 10,000 tokens), 2350 (138.5), and 334 (164.0) respectively (Fig. 1); and among ninety eight types of LA, NS, JLE, and CLE use 32, 28, 22 types respectively (Fig. 2).

First, in terms of the amount, neither the difference between NS and JLE (x^2=0.01, *ns*) nor that between NS and CLE (x^2=5.6, *ns*) was significant. European NNS' tendency to use fewer tokens of LA items, as pointed out by Altenberg & Tapper (1998) and Eia (2006), was not clearly supported here.

Second, in terms of variety, neither the difference between NS and JLE (x^2=2.06, *ns*) nor that between NS and CLE (x^2=1.74, *ns*) was significant. European NNS' tendency to use more varied LA types than NS was not clearly supported in the case of Asian NNS, who tend to use a narrower range of LA items.

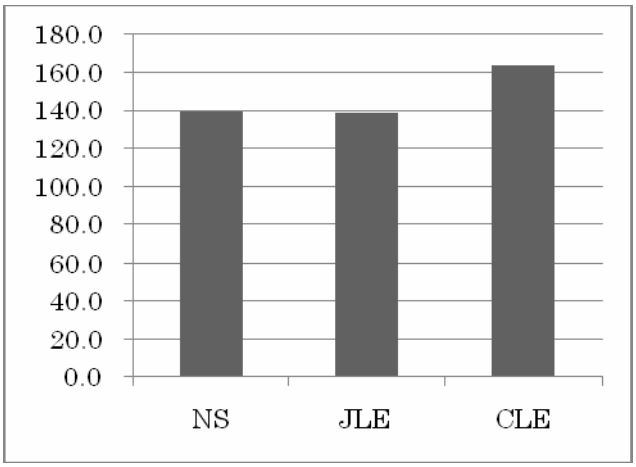

Figure 1. Tokens of LA items

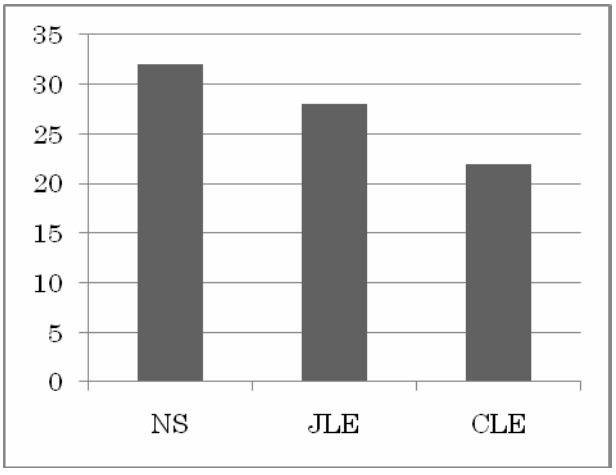

Figure 2. Types of LA items

It attracts our attention that the findings in the previous studies, which state that European NNS tend to use fewer LA items with greater variety, were not clearly supported here. One possible interpretation is that the writing conditions for the corpora used in previous studies were not strictly controlled hence the results might have been biased. Other interpretations relate to the difference of NNS' L1s and to that in the types of English education they received. For European NNS whose L1s have morphologically and functionally similar LA systems, using varied English LA

items seems relatively easy, but not for Asian NNS whose L1s have no compatible LA systems. Also, NNS in Japan and China, where English is taught primarily as an academic language, have been generally instructed to make their essays as logical as possible, and they do not necessarily avoid using LA in essays like their European counterparts.

As the previous studies and the current study are based on different corpora and different research designs, it might not be appropriate to directly compare their findings. However, it is important that our analysis showed the possibility that the NS/NNS gap might not be as essential as previously suggested in terms of the LA usage.

LA types

As shown in the following table, the top five functional types are concessive, addictive, meta-textual, listing, and contrastive in all cases of NS, JLE, and CLE. The overall tendencies concerning the choice of the LA functional types are largely identical among all three writer-groups (Table 4).

Now, let us focus on the top five functional types. The figure below shows how often NS, JLE, and CLE use the five major LA functional types. JLE and CLE seem to use concessive and addictive types of LA more often than NS, but the NS/NNS differences are only significant in the case of the addictive type (NS/JLE: x^2=7.52; NS/CLE: x^2=8.14).

Table 4. Functional types of LA

Functions	Raw Freq.			Adjusted Freq.		
	NS	JLE	CLE	NS	JLE	CLE
addictive	166	953	127	44.66	56.17	62.36
concessive	189	898	127	50.84	52.93	62.36
contrastive	27	26	10	7.26	1.53	4.91
listing	43	136	19	11.57	8.02	9.33
meta-textual	70	317	47	18.83	18.69	23.08
resultive	5	7	1	1.35	0.41	0.49
summative	7	2	0	1.88	0.12	0.00
time	10	11	3	2.69	0.65	1.47

Also, we can gain much more than what we have learned from books. (CLE)

But I do not think that I want my parents to pay for money of these things, ***too***. (JLE_M)

Furthermore, teenagers do not have a formed standard, they do not have a clear thought on what is good or not. (CLE)

The cumulative adjusted frequencies of the top three functional types are 114.33 for NS, 127.79 for JLE, and 147.80 for CLE respectively. Asian NNS tend to be more exclusively dependent on a few major functional types of LA, which suggests that they are not sufficiently aware of the possible variety of LA.

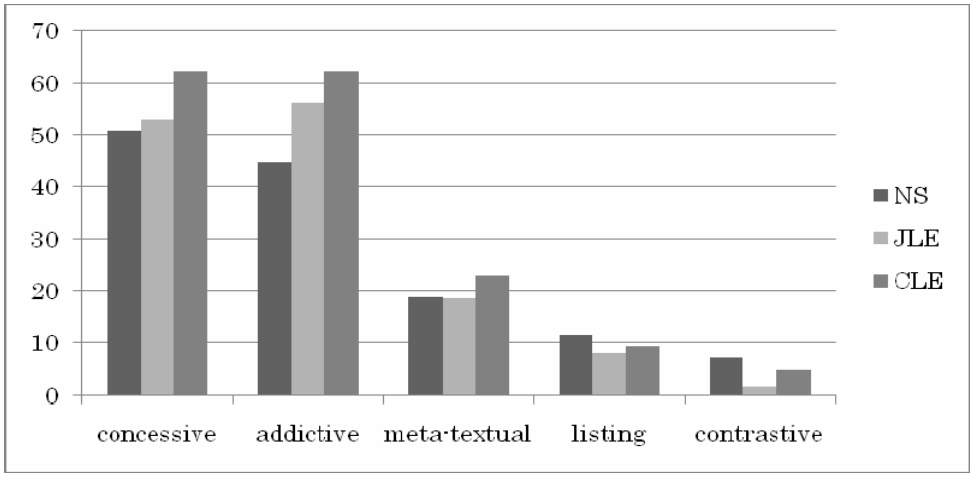

Figure 3. Adjusted frequencies of five major LA functional types

LA items

Next, focusing on the top twenty LA items, we identified LA items statistically overused or underused by JLE and CLE. In Table 5, the symbols (+) or (-) indicate that the adjusted frequency for NNS is higher or lower than that for NS respectively. The asterisks attached to the symbol show that the differences between frequencies are significant at the five percent level. Also, the Dif column summarizes the over/underuse tendency by NNS: "+ +" suggests that the listed LA item is overused both by JLE and by CLE; "+ -" suggests that it is overused by JLE while underused by CLE, for instance.

Table 5. Adjusted frequency of the top twenty LA items

Item	Function	NS	JLE		CLE		Dif
again	addictive	1.35	0.47	(-)	0.00	(-)	- -
also	addictive	29.05	32.48	(+)	48.12	(+) *	+ +
furthermore	addictive	2.15	2.06	(-)	0.98	(-)	- -
moreover	addictive	1.08	5.78	(+) *	4.91	(+) *	+ +
too	addictive	9.68	15.21	(+) *	8.35	(-)	+ -
anyway	concessive	0.54	0.94	(+)	0.98	(+)	+ +
however	concessive	12.64	12.79	(+)	10.31	(-)	+ -
only	concessive	21.52	32.12	(+) *	32.90	(+) *	+ +
still	concessive	8.07	1.65	(-)	8.35	(+)	- +
though	concessive	3.23	4.24	(+)	8.35	(+) *	+ +
yet	concessive	3.50	0.88	(-) *	0.98	(-)	- -
instead	contrastive	2.15	0.65	(-) *	0.00	(-)	- -
rather	contrastive	4.57	0.88	(-)	1.47	(-)	- -
finally	listing	5.92	3.95	(-)	0.98	(-) *	- -
lastly	listing	1.08	1.00	(-)	0.00	(-)	- -
next	listing	4.57	3.07	(-)	0.98	(-)	- -
indeed	meta-textual	2.42	1.00	(-)	0.49	(-)	- -
now	meta-textual	6.73	13.26	(+) *	12.27	(+)	+ +
well	meta-textual	8.88	4.13	(-) *	10.31	(+)	- +
eventually	time	2.42	0.47	(-) *	0.00	(-)	- -

There are several noteworthy findings. First, the difference between NS and NNS is not as clear as suggested in previous studies. The number of statistically significant overused or underused items is eight of twenty (forty percent) for JLE, and only five (twenty percent) for CLE. The difference is not observed in most of the LA types.

Second, the overall over/underused trends of JLE and CLE are largely identical. Agreements shown as "+ +" or " - - " in the right column are observed in as many as sixteen of twenty LA items (eighty percent), although the differences in the frequencies between NS and NNS are not always significant.

Finally, JLE and CLE overuse the LA items concerning the introduction of additional information (*also*, *moreover*), and those concerning the intensification of meaning

(*anyway*, *only*, *now*). Among these, *moreover* and *only* are significantly overused by both JLE and CLE.

> It is difficult to learn these things ***only*** in college because the relation between students contains little social responsibility. (JLE_S)
>
> ***Moreover***, as college students, aside from academic knowledge, we should gain more experience through those part-time job. (CLE)

While they underuse the LA items related to the introduction of parallel information (*instead*, *rather*) and those related to sequences of presented information (*finally*, *lastly*, *eventually*), Asian NNS seem to prefer adding new information cumulatively to old information rather than presenting or listing several pieces of information simultaneously.

Relation between writer groups and LA use

Finally, we conducted correspondence analysis to see how seven kinds of writer-groups (NS, JLE, CLE; JLE_Lower, JLE_Middle, JLE_Semi-upper, JLE_upper), which are set as item 1, are classified, and which of the twenty major LA items, which are set as item 2, characterize different writer groups. The result of the analysis is shown in Table 6.

Table 6. Result of correspondence analysis

	Dim. 1	Dim. 2	Dim. 3	Dim. 4	Dim. 5
Eigenvalue	0.0429	0.0139	0.0039	0.0034	0.0013
Contribution	65.58	21.23	5.94	5.2	2.05
Cum. Contribution	65.58	86.81	92.75	97.95	100

As the dimensions 1 (z_1) and 2 (z_2) cover 86.81% of the entire data distribution, we created scatter plots by making dimensions 1 and 2 the horizontal and vertical axes respectively, on which the individual category data of items 1 (writer-groups) and 2 (LA items) are shown. There are several noteworthy findings here. First, NS, JLE, and CLE make their own clusters as shown in Fig. 4a. Although in terms of quantity, variety, functional types, and over/underused LA items, the gap between NS and Asian NNS has not been observed so clearly, the result of correspondence analysis suggests that each of the three writer groups still has its own LA usage pattern, even with the commonest LA items.

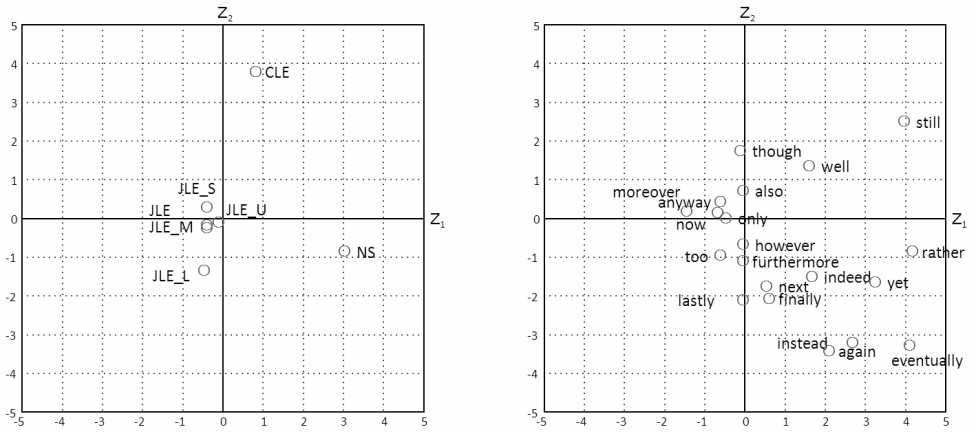

Fig. 4. a) Item 1 scatter plot, b) Item 2 scatter plot

Second, the Z_1 axis, which explains 65.58% of the data distribution, distinguishes between NS and other writer-groups; and the Z_2 axis, which explains 21.23% of the same, distinguishes between CLE and others. This suggests that the difference between NS and two NNS groups is much larger than that between two NNS groups. It is important that an essential boundary is observed between NS and Asian NNS in terms of LA use. Also, focusing on the Z_1 axis, we realize that CLE is somewhat closer to NS than is JLE.

Third, the L2 proficiency of JLE has a relatively smaller influence over the LA usage pattern. Although some difference is observed on the Z_2 axis between JLE at a lower proficiency level and JLE at other levels, JLE at middle, semi-upper, and upper levels seem to use LA in more or less identical ways. Also, all the JLE subgroups are positioned vertically in a row and show little tendency to move rightward, namely, toward NS, which suggests that JLE's deviant use of LA is not corrected even if their L2 proficiency level increases. This could be at least partly attributed to the nature of English education given to JLE. As is often pointed out, the amount of target language JLE read or listen to is highly limited compared with other Asian EFL learners. According to Tono (2008), the number of the words (types) appearing in junior high school textbooks in Japan is half that in Korea and a third of that in China. Lack of L2 input, which is a baseline for judging naturalness in L2 use, generally deprives JLE of the opportunities to realize how deviant their L2 use is. This seems to cause JLE to linger on a rather peculiar usage of LA, even if it is not clearly ungrammatical.

Fourth, as shown in Fig. 4b, each of the three writer groups uses a particular set of LA items: NS are characterized most typically by LA items such as *yet*, *rather*, and *indeed*; while JLE by *only*, *however*, and *too*.

> The income may be low for the students, ***yet*** it yields a double bonus with respect to the money they actually earn, and the money they save from not spending while working. (NS)
>
> It is not the question of whether we should accommodate smokers or not, but ***rather*** a question of how we should accommodate them. (NS)
>
> If they wish to accept the harm that smoking related illnesses cause, it is ***indeed*** their right. (NS)
>
> Some people say that if we separate people by making smoking section and non-smoking section, there will be no problem. ***However***, I think that it is not inefficient way for restaurants. (JLE_U)

There are no LA items clearly characterizing CLE, but a set of words such as *though*, *well*, and *still*, whose category scores are relatively high in Z_2, seem to have some affinity to CLE. This kind of information will offer NNS a clue to understand how NS use LA in their writing and how their own LA uses differ from the standard.

Conclusion

In the current study, we observed how NS, JLE, and CLE use LA in argumentative essays. Our corpus-based investigation has yielded several essential facts concerning the frequency and usage of LA. Our findings can be summarized as follows:

1) Do Asian NNS use fewer but more varied LA compared to NS?

   Previous studies have suggested that European NNS tend to use LA items less often but with greater variety compared to NS. However, in our data of Asian NNS, no significant difference emerged in either the tokens or the types of LA items. This discrepancy would be explained by the differences in (i) the designs of learner corpora used for analyses, (ii) L1s of the NNS, and (iii) the types of English education given to them. Our analysis suggested the possibility that the NS/NNS gap in LA use might not be as essential as suggested before.

2) Which functional type of LA is overused or underused by Asian NNS?

   The frequently used functional types are concessive, addictive, meta-textual, listing, and contrastive in all cases of NS, JLE, and CLE. The difference between NS and Asian NNS is not necessarily large, but the latter tend to overuse addictive types. Also, Asian NNS tend to be dependent on a limited number of LA types.

3) Which item of LA is overused or underused by Asian NNS?

   The gap in frequencies between NS and Asian NNS is not statistically significant in most of the top twenty LA items. However, Asian NNS have a tendency to overuse LA items concerning the introduction of additional information (*also*, *moreover*), and those concerning the intensification of meaning (*anyway*, *only*, *now*), while often underusing LA items concerning the introduction of parallel information (*instead*, *rather*) and those concerning the sequential introduction of information (*finally*, *lastly*, *eventually*).

4) To what extent are JLE at varied L2 proficiency levels, CLE, and NS close to each other in terms of the usage of LA? JLE, CLE, and NS are positioned in different clusters, suggesting that these three groups of writers use LA in their own ways. JLE is positioned the farthest from NS and it was proven that the increase of L2 proficiency of JLE does not necessarily lead to accession to NS' standard usage pattern of LA. Also, it emerged that the so-called native-likeness is closely related to LA items such as *yet*, *rather*, and *indeed*.

It was suggested that the increasing L2 proficiency does not directly mean an NS-like natural use of LA in the target language; therefore, these findings need to be appropriately demonstrated to L2 learners. What is pedagogically important is to make learners aware of such NS/NNS gap. Unlike in the case of children's L1 acquisition, where a great amount of language input plays a decisive role, explicit teaching and awareness-raising tasks would be vital for adult L2 learners (Ellis, 1991). ICT can undoubtedly contribute, not only to a data-based language study, but also to the design of a new language curriculum and teaching materials based on the findings obtained from corpus analysis.

Author Note

This is an entirely revised version of a paper read at the 2009 ICTATLL (Information and Communications Technology in Analysis, Teaching and Learning of Languages) Conference held at the University of the Peloponnese, in Corinth, Greece, 10-11 September 2009.

References

Aarts, J. (1991). Intuition-based and observation-based grammars. In K. Aijmer & B. Altenberg (Eds.), *English corpus linguistics: Studies in honour of Jan Svartvik* (pp. 44-62). London, UK: Longman.

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. 55-77). Amsterdam, The Netherlands: Benjamins.

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, UK: Addison Wesley Longman.

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

Biber, D., & Reppen, R. (1998). Comparing native and learner perspectives on English grammar: A study of complement clauses. In S. Granger (Ed.), *Learner English on computer* (pp. 145-158). London, UK: Addison Wesley Longman.

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

De Cock, S., Granger, S., Leech, G., & McEnery, T. (1998). An automated approach to the phrasicon of EFL learners. In S. Granger (Ed.), *Learner English on computer* (pp. 67-79). London, UK: Addison Wesley Longman.

Eia, A. (2006). The use of linking adverbials in Norwegian advanced learners' written English. Unpublished Master's Thesis. Oslo, Norway: Oslo University.

Ellis, R. (1991). *Second language acquisition & language pedagogy*. Bristol, UK: Multilingual Matters Ltd.

Granger, S. (1998). The computerized learner corpus: A versatile new source of data for SLA research. In S. Granger (Ed.), *Learner English on computer* (pp. 3-18). London, UK: Addison Wesley Longman.

Granger, S. (2002). A bird's-eye view of computer 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, The Netherlands: Benjamins.

Hasselgren, A. (2002). Learner corpora and language testing: Small words as markers of leaner proficiency. In S. Granger, J. Hung, & S. Petch-Tyson (Eds.), *Computer learner corpora, second language acquisition and foreign language teaching* (pp. 143-174). Amsterdam, The Netherlands: Benjamins.

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, The Netherlands: Benjamins.

Ishikawa, S. (2008). *Eigo ko-pasu to gengo kyouiku*. Tokyo, Japan: Taishukan Shoten. [English corpus linguistics for language education].

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

Lorenz, G. (1998). Overstatement in advanced learners' writing: Stylistic aspects of adjective intensification. In S. Granger (Ed.), *Learner English on computer* (pp. 53-66). London, UK: Addison Wesley Longman.

Narita, M., Sato, C., & Sugiura, M. (2003). The use of linking adverbials in the English essay writing of Japanese EFL learners. In M. Murata, S. Yamada, & Y. Tono (Eds.), *ASIALEX '03 Tokyo Proceedings* (pp.440-444). The Asian Association for Lexicography.

Petch-Tyson, S. (1998). Writer/reader visibility in EFL written discourse. In S. Granger (Ed.), *Learner English on computer* (pp. 107-118). London, UK: Addison Wesley Longman.

Quirk, R., & Greenbaum, S. (1973). *A university grammar of English*. Harlow, UK: Longman.

Ringbom, H. (1998). Vocabulary frequencies in advanced learner English: A cross-linguistic approach. In S. Granger (Ed.), *Learner English on computer* (pp. 41-52). London, UK: Addison Wesley Longman.

Sinclair, J. (2005). *Collins COBUILD English grammar*. Glasgow, UK: Harper Collins (2nd ed.).

Svartvik, J. (Ed.). (1990). *The London-Lund corpus of spoken English: Description and research*. Lund, Sweden: Lund University Press.

Tono, Y. (2008). Ajia kakkoku to nihon no eigo kyoukasyo hikaku. [Comparison of English Textbooks in Japan and Asian Countries]. Retrieved March, 20, 2010, from http:// www.kantei.go.jp/jp/singi/kyouiku_kondan/kaisai/dai3/2seku/2s-siryou3.pdf

Appendix

Table 7. Raw and adjusted frequencies of all LA items appearing in the corpus

Item	Function	Type	Raw Freq.			Adjusted Freq.			
			NS	JLE	CLE	NS	JLE	CLE	Av.
also	addictive	W	108	551	98	29.05	32.48	48.12	36.55
only	concessive	W	80	545	67	21.52	32.12	32.90	28.85
however	concessive	W	47	217	21	12.64	12.79	10.31	11.92
too	addictive	W	36	258	17	9.68	15.21	8.35	11.08
now	meta-textual	W	25	225	25	6.73	13.26	12.27	10.75
well	meta-textual	W	33	70	21	8.88	4.13	10.31	7.77
still	concessive	W	30	28	17	8.07	1.65	8.35	6.02
though	concessive	W	12	72	17	3.23	4.24	8.35	5.27
moreover	addictive	W	4	98	10	1.08	5.78	4.91	3.92
finally	listing	W	22	67	2	5.92	3.95	0.98	3.62
next	listing	W	17	52	2	4.57	3.07	0.98	2.87
firstly	listing	W	0	0	15	0.00	0.00	7.36	2.45
rather	contrastive	W	17	15	3	4.57	0.88	1.47	2.31
instead	contrastive	W	8	11	7	2.15	0.65	3.44	2.08
yet	concessive	W	13	15	2	3.50	0.88	0.98	1.79
furthermore	addictive	W	8	35	2	2.15	2.06	0.98	1.73
indeed	meta-textual	W	9	17	1	2.42	1.00	0.49	1.30
eventually	time	W	9	8	0	2.42	0.47	0.00	0.96
anyway	concessive	W	2	16	2	0.54	0.94	0.98	0.82
lastly	listing	W	4	17	0	1.08	1.00	0.00	0.69
again	addictive	W	5	8	0	1.35	0.47	0.00	0.61
overall	summative	W	5	2	0	1.35	0.12	0.00	0.49
equally	addictive	W	4	3	0	1.08	0.18	0.00	0.42
nevertheless	concessive	W	1	5	1	0.27	0.29	0.49	0.35
namely	meta-textual	W	3	4	0	0.81	0.24	0.00	0.35
hence	resultative	W	2	0	1	0.54	0.00	0.49	0.34
meanwhile	time	W	0	0	2	0.00	0.00	0.98	0.33
consequently	resultative	W	3	1	0	0.81	0.06	0.00	0.29
nonetheless	concessive	W	3	0	0	0.81	0.00	0.00	0.27
afterwards	time	W	1	0	1	0.27	0.00	0.49	0.25
alternatively	contrastive	W	2	0	0	0.54	0.00	0.00	0.18

Item	Function	Type	Raw Freq.			Adjusted Freq.			
			NS	JLE	CLE	NS	JLE	CLE	Av.
altogether	summative	W	2	0	0	0.54	0.00	0.00	0.18
accordingly	resultative	W	0	6	0	0.00	0.35	0.00	0.12
admittedly	concessive	W	1	0	0	0.27	0.00	0.00	0.09
likewise	addictive	W	1	0	0	0.27	0.00	0.00	0.09
originally	time	W	0	3	0	0.00	0.18	0.00	0.06
incidentally	meta-textual	W	0	1	0	0.00	0.06	0.00	0.02

Table 8. LA items which are not observed in the corpus

Item (Function, Type)
above all (addictive: P), after all (concessive: P), after that (time: P), all in all (summative: P), all the same (concessive: P), anyhow (concessive: W), as a consequence (resultative: P), as a result (resultative: P), as well (addictive: P), at any rate (concessive: P), at the same time (concessive: P), by comparison (contrastive: P), by contrast (contrastive: P), by the way (meta-textual: P), correspondingly (addictive: W), first of all (listing: P), for a start (listing: P), for all that (concessive: P), for example (meta-textual: P), for instance (meta-textual: P), for one thing (listing: P), in addition (addictive: P), in any case (concessive: P), in any event (concessive: P), in comparison (contrastive: P), in conclusion (summative: P), in contrast (contrastive: P), in other words (meta-textual: P), in particular (addictive: P), in short (summative: P), in spite of that (concessive: P), in sum (summative: P), in summary (summative: P), in that case (inference: P), in the first place (listing: P), in the meantime (time: P), in the second place (listing: P), last of all (listing: P), meantime (time: W), mind you (concessive: P), more accurately (contrastive: P), more precisely (contrastive: P), on the contrary (contrastive: P), on the one hand (listing: P), on top of it all (addictive: P), or rather (meta-textual: P), so to speak (meta-textual: P), subsequently (time: W), that is (meta-textual: P), that is to say (meta-textual: P), that said (concessive: P), then again (contrastive: P), to begin with (listing: P), to cap it all (addictive: P), to conclude (summative: P), to crown it all (addictive: P), to put it another way (meta-textual: P), to start with (listing: P), to sum up (summative: P), to summaries (summative: P), what is more (addictive: P),