

# Investigating Chinese College Learners' Use of Frequency Adverbs: A Corpus-based Approach\*

Guobing Liu

Faculty of International Studies/Corpus Research and Application Centre, Henan Normal University, Xinxiang, China

**Abstract**—The present study aims at finding out the differences between Chinese college English learners and the native speakers with regard to the use of frequency adverbs. This paper, based on corpus linguistics, is intended to examine how Chinese college English learners use the 20 most frequently used frequency adverbs, or TTFAs, the top twenty frequency adverbs in their written and spoken English and how they use these TTFAs differently from the way native speakers use them as recorded in the native speakers' corpora. From the research result, we can find that Chinese college learners tend to overuse and underuse certain TTFAs in their speech and writing. The overusing tendency is slightly stronger than the underusing tendency in both speaking and writing. The use of TTFAs is found to be problematic for Chinese college English learners. Compared with the native speakers, Chinese college learners seem to have a multifold problem. From the perspective of language learning, the research findings shed light on the problems for the learners in their vocabulary acquisition and EFL learning. Chinese college learners should try to enlarge their vocabulary and vary their choice of TTFAs in their writing and speaking. At the same time, they also should develop a register-awareness in choices of frequency adverbs.

**Index Terms**—frequency adverb, corpus, TTFAs, use

## I. INTRODUCTION

Frequency adverbs are a category of adverbs which can tell us how often something happens, happened or will happen (Wen and Ting, 2003). Spoken and written English texts contain large quantities of frequency adverbs such as "often", "never", "always", "sometimes" and etc. A major part of English adverbs as they belong to, frequency adverbs play an extremely important role in both oral and written communication. Through using frequency adverbs, the language user can express how often something is, was or will be the case (Biber et al., 2000). More usually, however, language users are concerned with frequency adverbs with respect to a specified or implied span of time and refer them to "the rate of occurrences per unit of time" with inexplicit numerical values (Leech & Svartvik, 1994).

A number of studies (such as Wen and Ting, 2003; Odlin, 1989; 2001; Leech et al., 2001) have shown that the use of frequency adverbs is problematic for language users, in particular for foreign language learners. However, these studies did not compare Chinese college English learners' use of frequency adverbs between different registers such as speech and writing, nor did they make a systematic comparison between Chinese college English learners and the native speakers with regard to the use of frequency adverbs. This paper, based on corpus linguistics, is intended to examine how Chinese college English learners use the 20 most frequently used frequency adverbs, or TTFAs, the top twenty frequency adverbs (Leech, et al. 2001; cited in Wen and Ting, 2003) in their written and spoken English and how they use these TTFAs differently from the way native speakers use them as recorded in the native speakers' corpora. The research questions addressed in this study are as follows:

- (1) Compared with the native speakers, do Chinese college English learners overuse or underuse frequency adverbs?
- (2) Do they overuse or underuse frequency adverbs differently across spoken-register and written-register?
- (3) Do they demonstrate a similar pattern of writing-speaking difference as the native speakers in the use of frequency adverbs? If not, what are the differences?
- (4) For such differences, can we find some reasons from different aspects such as culture, society and education?

## II. DATA AND METHODOLOGY

The computer learner corpora used in this study included a written English corpus, 474,511 words of essay writing by college English learners from the Chinese Learner English Corpus (CLEC) developed by Gui and Yang (2003), and the College Learners' Spoken English Corpus (COLSEC), which funded by the Chinese National Social Science Research Foundation and completed in the early 2004. CLEC and COLSEC were designed to be sister corpora. The native speaker control corpora used in this study are Louvain Corpus of Native English Essays (LOCNESS), a sub-corpus of International Corpus of Learner English (ICLE), which contains 181,879-word essays written by British and American

---

\* This research was supported in part by the Ministry of Education in China under the Project of Humanities & Social Sciences (Project No.: 12YJC740062) and Henan Planning Office of Philosophy & Social Sciences with the Project of Philosophy and Social Sciences (Project No.: 2013CYY025).

students and a spoken English corpus, LONDON-LUND Corpus. Table I provides detailed information of the four corpora used in this study. We use WordSmith Tools (Scot, 1998) and PowerConc (Liang et al. 2013) to find the occurrence frequencies of each TFA in the four corpora respectively and then to compare them. Since the sizes of the four corpora are different, all the raw counts are computed into normalized frequencies (occurrences per 1,000,000 words).

TABLE I.  
DESCRIPTIVE DATA OF LEARNERS' CORPORA AND NATIVE SPEAKER'S CORPORA

Type of corpus		Size of the corpus	Total words
The learner corpora	CLEC (ST3, ST4)	474,511	1,182,972
	COLSEC	708,461	
The control corpora	LOCNESS	181,879	715,142
	LONDON-LUND	533,245	

### III. FINDINGS AND RESULTS

#### A. Normalized Frequencies of TFAs in Learner Corpora

The first research question of this study concerns the general tendency of the use of TFAs by Chinese college English learners. Figure 1 brings out a significant difference between Chinese college students and the native speakers with regard to the use of TFAs. All together, we studied 20 frequency adverbs in this study. There are only 3 frequency adverbs that Chinese learners and the native speakers use similarly. The rate of similarity is 15%. However, Chinese college learners use 85% of the TFAs differently with the native speakers. Among them, overuse occupies 45%, underuse for 40% respectively. From the table II we can see that in the learner corpora, there are 8103 occurrences of the TFAs per million words while in the control corpora, there are only 4314 occurrences per million tokens. Chinese college learners use 3789 frequency adverbs more than the native speakers. Table II also shows us that Chinese college English learners not only overuse (9) more TFAs than they underuse (8) them, but the overusing tendency is much stronger than the underusing tendency. The difference in average normalized frequency between the two corpora is far greater in the overused TFAs (473 occurrences) than in the underused ones (57 occurrences).

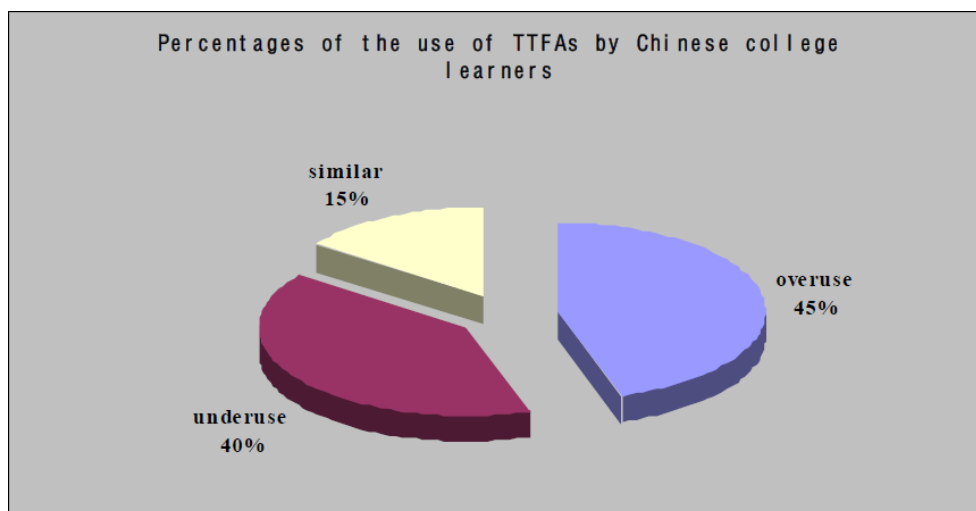


Figure 1: Pie chart of TFAs used by Chinese college learners

TABLE II.  
AN OVERALL COMPARISON BETWEEN THE LEARNERS' CORPORA AND THE CONTROL CORPORA

	TTFAS	The learner corpora	The control corpora	Difference	Tendencies	Degree of overuse or underuse
1	always	1910	799	1111	overuse	9 overused TTFAs Average normalized frequency difference 473
2	often	1723	470	1253	overuse	
3	sometimes	1094	116	978	overuse	
4	usually	727	161	566	overuse	
5	never	666	726	-60	underuse	
6	ever	565	605	-40	underuse	
7	once	531	509	22	overuse	
8	constantly	206	117	89	overuse	
9	frequently	199	24	175	overuse	
10	generally	103	114	-11	underuse	
11	hardly	72	83	-11	underuse	8 underused TTFAs Average normalized frequency difference 57
12	in general	64	42	22	overuse	
13	regularly	61	24	37	overuse	
14	mostly	56	60	-4	similar	
15	no longer	35	246	-211	underuse	
16	twice	28	73	-45	underuse	
17	increasingly	23	77	-54	underuse	
18	normally	17	41	-24	underuse	
19	occasionally	14	9	5	similar	
20	rarely	9	18	-9	similar	
Average normalized frequency		8103	4314	3789		

B. Comparison of Chinese Learners' Use of TTFAs with the Native Speakers' in Their Writing

Table III brings us the fact that Chinese learners tend to use more TTFAs in their writing. In the CLEC, the average normalized frequency of TTFAs is 267. However, there are only 155 occurrences of TTFAs in the LOCNESS. Chinese college learners use 112 frequency adverbs more than the native speakers. The difference is significant. But through the close examination of the normalized frequency difference in each TFA in the CLEC and LOCNESS, we can find both overuse and underuse of TTFAs exist in Chinese college learners' English writing. Chinese college learners not only overuse more TTFAs (11 TTFAs) than they underuse them (5 TTFAs), but the overusing tendency is much stronger than the underusing tendency because the difference in average normalized frequency in the overused TTFAs (237 occurrences per million words) between CLEC and LOCNESS is much greater than that in the underused TTFAs (71 occurrences per million words).

TABLE III.  
A COMPARISON BETWEEN CLEC AND LOCNESS

	TTFAs	CLEC	LOCNESS	Difference	Degree of overuse or underuse
1	often	1271	423	848	11 overused TTFAs Average normalized frequency difference 237
2	always	1206	506	700	
3	sometimes	398	93	305	
4	usually	432	137	295	
5	frequently	169	22	147	
6	ever	398	269	129	
7	constantly	202	115	87	
8	never	485	445	40	
9	once	417	396	21	
10	in general	57	38	19	
11	regularly	27	16	11	
12	occasionally	13	5	8	Similar 4
13	hardly	65	66	-1	
14	mostly	48	49	-1	
15	rarely	8	16	-8	
16	normally	6	22	-16	
17	generally	72	110	-38	5 underused TTFAs Average normalized frequency difference 71
18	twice	21	60	-39	
19	increasingly	19	77	-58	
20	no longer	27	231	-204	
Average normalized frequency		267	155	112	

C. Comparison of Chinese Learners' Use of TTFAs with the Native Speakers' in Their Speaking

Table IV shows how Chinese college learners overuse and underuse TTFAs in their speech. In the learners' spoken corpus COLSEC, there are 138 occurrences of TTFAs per million words, but in the native speakers' spoken corpus LUND there are only 61 occurrences per million words. Chinese college learners used 77 TTFAs more than the native speakers in one-million-word speaking. The difference exists, but not significant as that in their writing. Chinese college learners overuse 7 TTFAs and underuse 3 ones. There are 10 TTFAs with similar frequencies to the native speakers. As

in their writing, Chinese college learners tend to overuse more TTFAs than they underuse them. Overusing tendency (average normalized frequency difference is 263) is slightly stronger than the underusing tendency (average normalized frequency difference is 93).

TABLE IV.  
A COMPARISON BETWEEN COLSEC AND LUND

	TTFAs	COLSEC	LUND	Difference	Degree of overuse or underuse
1	sometimes	696	23	673	7 overused TTFAs Average normalized frequency difference 263
2	always	704	293	411	
3	often	452	47	405	
4	usually	295	24	271	
5	frequently	30	2	28	
6	generally	31	4	27	
7	regularly	34	8	26	
8	increasingly	4	0	4	Similar 10
9	in general	7	4	3	
10	constantly	4	2	2	
11	once	114	113	1	
12	rarely	1	2	-1	
13	mostly	8	11	-3	
14	occasionally	1	4	-3	
15	twice	7	13	-6	
16	no longer	8	15	-7	
17	normally	11	19	-8	
18	hardly	7	17	-10	3 underused TTFAs Average normalized frequency difference 93
19	never	181	281	-100	
20	ever	167	336	-169	
Average normalized frequency		138	61	77	

#### D. Comparison of Chinese College Learners' Use of TTFAs in Their Writing and Speaking

Table V is a summary of the difference between the use of TTFAs by Chinese college English learners and the native speakers. In CLEC, Chinese learners overuse 11 TTFAs per million words. The average normalized frequency difference is 237. While in COLSEC, though Chinese college learners overuse 7 TTFAs per million tokens and the number is less than that in CLEC, the average normalized frequency difference is 263. It indicates that the overusing tendency is much stronger in speech than in writing. Similarly, in terms of the underusing tendency, Chinese college learners underuse 5 TTFAs in one-million-word writing and only 3 TTFAs in the same length speaking. But the average normalized frequency difference in COLSEC is 93. The underusing tendency of Chinese college learners in speaking is stronger than that in writing. In CLEC, Chinese college learners use 4 TTFAs similarly with the native speakers. While in COLSEC, there are 10 TTFAs with the same occurrences. But in COLSEC, the average normalized frequency difference (-18) is much bigger than that in CLEC. Generally speaking, Chinese college learners differ more from the native speakers in speaking than they do in writing with regard to the use of TTFAs.

TABLE V.  
A COMPARISON OF TWO SETS OF FREQUENCY DIFFERENCES

		CLEC-LOCNESS	COLSEC-LUND
Overuse	Number of TTFAs	11	7
	Average normalized frequency difference	237	263
Underuse	Number of TTFAs	5	3
	Average normalized frequency difference	71	93
Similar	Number of TTFAs	4	10
	Average normalized frequency difference	-2	-18

#### E. Writing-speaking Differences of TTFA Use in Chinese Learners' Corpora and the Native Speakers' Corpora

Table VI is a detailed description of the writing-speaking difference of TTFA use both in Chinese learners' corpora and the native speakers' corpora. We can find that the difference in learners' corpora (129) and the native speakers' corpora (50) are all positive. That is to say, both Chinese college learners and the native speakers use more TTFAs in writing than they do in their speaking. However, the average writing-speaking difference of the TTFA use by Chinese college learners is much bigger than that by the native speakers (129>50). We also can find the much more visualized difference from figure 2. These two groups show us two totally different writing-speaking difference patterns.

TABLE VI.  
A COMPARISON OF REGISTER DIFFERENCES BETWEEN THE LEARNERS' CORPORA AND THE NATIVE SPEAKERS' CORPORA

	TTFAs	Difference LOCNESS-LUND		TTFAs	Difference CLEC-COLSEC
1	often	376	1	never	819
2	once	283	2	often	502
3	no longer	216	3	usually	304
4	always	213	4	once	303
5	never	164	5	normally	231
6	usually	113	6	sometimes	198
7	constantly	113	7	rarely	139
8	generally	106	8	always	137
9	increasingly	77	9	constantly	58
10	sometimes	70	10	hardly	50
11	hardly	49	11	mostly	41
12	twice	47	12	increasingly	40
13	mostly	38	13	generally	19
14	in general	34	14	occasionally	15
15	frequently	20	15	regularly	14
16	rarely	14	16	in general	12
17	regularly	8	17	frequently	7
18	normally	3	18	ever	-5
19	occasionally	1	19	twice	-7
20	ever	-67	20	no longer	-298
Average difference		50	Average difference		129

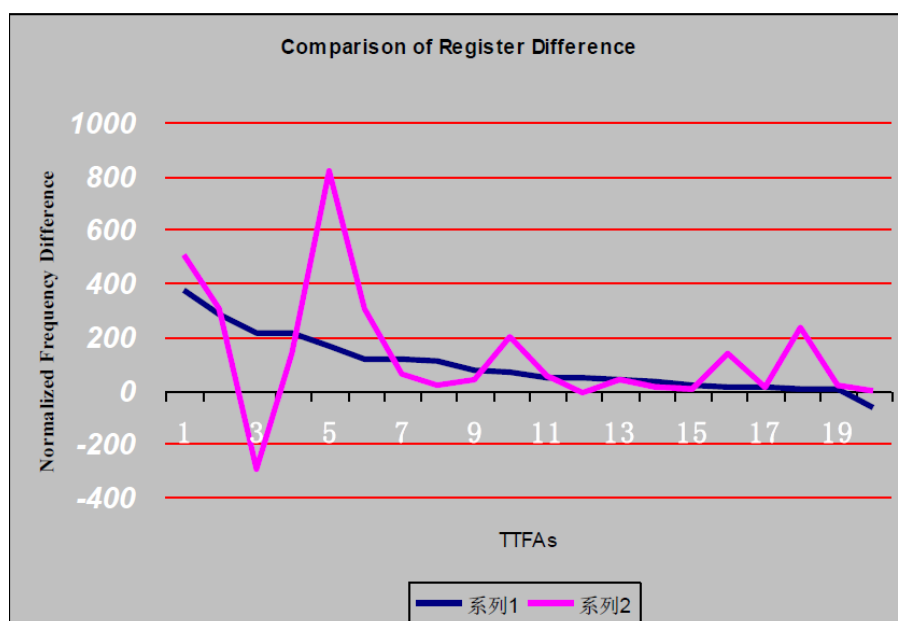


Figure 2: Register differences of Chinese learners and the native speakers

Notes: 系列 1 indicates learners' writing-speaking differences of each TTFa. 系列 2 stands for the writing-speaking difference of each TTFa of the native speakers. On X axis, from 1 to 20, the sequence of TTFAs is often, once, no longer, always, never, usually, constantly, generally, increasingly, sometimes, hardly, twice, mostly, in general, frequently, rarely, regularly, normally, occasionally and ever.

F. Comparison of Register Difference of TTFAs Used by the Learners and the Native Speakers

Table VII is a comparison of the differences of TTFa use between Chinese college learners and the native speakers with regard to their registers. As it shows, most of the TTFAs are written-register sensitive. On the use of the TTFAs, both Chinese college learners and the native speakers think that most of them are written-register sensitive, only 1 for spoken-register sensitive and 3 for register neutral. However, a close examination of table VII shows the complexities of the picture. Chinese college learners have a significant register difference with the native speakers on the use of TTFAs. In the native speakers' corpora, "ever" is spoken-register sensitive; but for Chinese college learners, "no longer" is spoken-register sensitive; in the learners' corpora, "regularly", "normally", "occasionally" are written-register sensitive, but they are register neutral in the native speakers' corpora. These findings tell us that Chinese college learners still do not have clear register awareness in their choice of TTFAs.

TABLE VII.  
REGISTER DIFFERENCE OF TTFA USE IN LEARNERS' CORPORA AND THE NATIVE SPEAKERS' CORPORA

	Learners' corpora		Native speakers' corpora	
	Number	TTFAs	Number	TTFAs
Spoken-register	1	<u>no longer</u>	1	ever
Written-register sensitive	16	never, often, usually, once, <u>normally</u> , sometimes, rarely, always, constantly, hardly, mostly, increasingly, <u>generally</u> , occasionally, <u>regularly</u> , in general.	16	often, once, no longer, always, never, usually, constantly, generally, increasingly, sometimes, hardly, twice, mostly, in general, frequently, rarely.
Register neutral	3	<u>frequently</u> , <u>ever</u> , <u>twice</u>	3	Regularly, normally occasionally

#### IV. DISCUSSIONS AND CONCLUSION

From the above study, it is found that Chinese college learners tend to overuse and underuse certain TTFAs in their speech and writing. Generally speaking, the overusing tendency is slightly stronger than the underusing tendency in both speaking and writing (see Table II and Figure 1). The overusing tendency is more marked in Chinese learners' writing than in their speaking according to the normalized frequencies of TTFAs (see Table V). Chinese college English learners overuse certain TTFAs but underuse some of them (see Table III and Table IV). They demonstrate a totally different pattern of speaking-writing difference with the native speakers (see Figure 2). Chinese learners display a general tendency to use more TTFAs in their writing (see Table VI). In sum, the use of TTFAs is found to be problematic for Chinese college English learners. Compared with the native speakers, they seem to have a multifold problem.

It's very hard for us to find the exact reasons why there is a so big difference between Chinese college learners and the native speakers with regard to the use of TTFAs. We can only give some possible factors that affect Chinese learners' choice of TTFAs. Firstly, Chinese college learners have a much smaller vocabulary than the native speakers. They try to overuse what they are familiar with, and to underuse even do not use what they are not. Secondly, the reasons why some TTFAs are overused by Chinese learners maybe come down to the Chinese language transfer. There are many high frequency words in Chinese which have the similar meanings with certain TTFAs, such as "often", "always", "usually", "never" and etc. Comparatively, some words are few. So Chinese learners tend to overuse the former, but underuse the latter. Thirdly, Chinese learners seem to lack a register-awareness, which reflects in their tendency to use many written-register-sensitive words such as "no longer" in their speaking. Conversely, they use spoken-register-sensitive word "ever" as register neutral. Finally, Chinese college English learners lack a full understanding of the semantic properties of some frequency adverbs, especially those with similar meaning.

#### V. CONCLUSION

From the perspective of language learning, the research findings shed light on the problems for the Chinese College English learners in their vocabulary acquisition and EFL learning. To solve the problems, several suggestions have been provided: Firstly, Chinese college learners should learn to differentiate the choices of frequency adverbs and try to acquire these words in a direct manner. Secondly, Chinese college English learners should place more emphasis on the restrictions imposed on certain frequency adverbs and improve semantic understanding of them. Thirdly, exposure to a greater range of registers to improve register awareness has been suggested. Finally, Chinese college English Learners should be aware that learning the native English writing and speaking conventions is inextricable from learning to write and speak. Put it in a simple way, Chinese college English learners can try their best to vary the choice of TTFAs both in their writing and speaking. At the same time, they also should develop a register-awareness in choices of frequency adverbs in their speaking and writing.

#### ACKNOWLEDGMENT

The author wishes to thank all the team members of the corpus research group in Henan Normal University. They provided me a lot of help during writing this paper. And I also want to express my special thanks to the responsible officer of Henan Normal University, who gave me the financial support under the project of initial fund for PhD faculties' research (2013).

#### REFERENCES

- [1] Altenberg, B. & Granger, S. (2002). The grammatical and lexical patterning of MAKE in native and non-native student writing. *Applied Linguistics* 22: 173-194.
- [2] Altenberg, B. & Tapper, M. (1998). The Use of Adverbial Connectors in Advanced Swedish Learner' Written English. In Granger, S. (Ed.), *Learner English on Computer*. London and New York: Longman.
- [3] Biber, D., Conrad, S. & Reppen, R. (1998). *Corpus Linguistics*. London: Cambridge University Press.
- [4] Biber, D., Johansson, S., Leech, G., Conrad, S. and Finegan, F. (2000). *Grammar of Spoken and Written English*. Beijing: Foreign Language Teaching and Research Press, 558-559.

- [5] Ellis, R. (1985). *Understanding Second Language Acquisition*. Oxford: Oxford University Press.
- [6] Gass, S., & Selinker, L. (Ed.). (1983). *Language Transfer in Language Learning*. Rowley: Newbury House.
- [7] 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), 19-29.
- [8] Gui S. & Yang H. (2003). *Chinese Learners' English Corpus*. Shanghai: Shanghai Foreign Language Education Press.
- [9] Hustonm, S. (2002). *Corpora in applied linguistics*. Cambridge: Cambridge University Press.
- [10] Leech, G., Rayson, P. & Wilson, A. (2001). *Word frequencies in written and spoken English*. London: Pearson Education Limited.
- [11] Leech, G. & Svartvik, J. (1994). *A communicative grammar of English*. London: Longman.
- [12] Liang M. (2009). The Mini-Text and Its Applications in Foreign Language Teaching. *Computer-Assisted Foreign Language Education*, 21(3): 8-12.
- [13] Liang, M., Xu, J. & Jia, Y. (2013). PowerConc 1.0. Beijing: Beijing Foreign Studies University.
- [14] Liu, G. (2003). On the Use of Linking Adverbials by Chinese College English Learners. *Journal of Language Teaching and Research*, 4(1), 149-156.
- [15] Liu, G. (2012). The Multi-Dimensional Perspectives of Corpus Linguistics—Review of CLIC 2011, *Journal of Chinese Foreign Language Education*, 12(4) 45-50.
- [16] Nation, I.S.P. (1990). *Teaching and learning vocabulary*. Massachusetts: Newbury House.
- [17] Odlin, T. (1989). *Language transfer*. Cambridge: Cambridge University Press.
- [18] Odlin, T. (2001). *Language Transfer, Cross-linguistic Influence in Language Learning*. Shanghai: Shanghai Foreign Language Education Press.
- [19] Oshima, A., & Hogue, A. (1997). *Coherence: Introduction to Academic Writing*. White Plains, NY: Addison Wesley Longman.
- [20] Quirk, R., Greenbaum, S., Leech, G., and Svartvil, J. (1985). *A Grammar of Contemporary English*. London: Longman, 631-647.
- [21] Selinker, L. (1972). Interlanguage. in Richards, J.C. (ed.) *Error analysis perspectives on second language acquisition*. London: Longman.
- [22] Sinclair, J.M. & Renouf, A. (1998). A lexical syllabus for language learning. in R. Carter & McCarthy, M. (eds.) *Vocabulary and language teaching*. London: Longman.
- [23] Scott, M. (1998). *WordSmith Tools (3.0)*. Oxford: Oxford University Press.
- [24] Wen, Q. & Ting, T. (2003). A Corpus-based Analysis of the Use of Frequency Adverbs by Chinese University English Majors. Paper presented at the 2003 Shanghai International Conference on Corpus Linguistics, Shanghai.
- [25] Yang, H. (2002). *Introduction to Corpus Linguistics*. Shanghai: Shanghai Foreign Language Education Press.

**Guobing Liu** was born in Huaxian County of Henan, China. He received the master's degree in corpus linguistics in 2005. Now he is an associated professor and EFL teacher at the Faculty of International Studies, Henan Normal University. He studied in Chinese National Research Center for Foreign Language Education, Beijing Foreign Studies University from 2010 to 2013 and received the doctoral degree in corpus linguistics and computational linguistics there. In recent years, he published two books and more than twenty academic papers in the key journals both home and abroad. He has also finished 3 research projects. His academic interests include corpus linguistics and foreign language teaching.