

A Corpus-Based Study on Learners' Common Errors in Word Classes in Their Writings

Lixin Xia¹

Abstract

The paper reports a study on the errors in word classes made by Chinese college students in their writings. From the CLEC, all the errors tagged as [wd2] were collected, and then a general overview of the errors among 4 groups of college students was given. After that, the first 100 errors with the greatest frequency were sorted out from all the errors. They were further classified into 8 categories according to the wrongly used word classes. Based on the actual errors in the CLEC, possible causes of these errors were identified and analyzed. Finally, potential implications for English teaching and learning were discussed and suggestions were put forward.²

Key words: Errors, Word Classes, Learner Corpora, English Pedagogical Lexicography, English Language Teaching

1. Introduction

A word class is "a set of words that display the same formal properties, especially their inflections and distribution." (Kolln & Funk, 2008) It is similar to the more traditional term part of speech. However, "When linguists began to look closely at English grammatical structure in the 1940s and 1950s, they encountered so many problems of identification and definition that the term part of speech soon fell out of favour, word class being introduced instead. Word classes are equivalent to parts of speech, but defined according to strict linguistic criteria." (Crystal, 2003) As a result, the term word class is adopted in the present study instead of part of speech.

Lots of studies have been done on word classes, but few of them looked at them from the perspective of second language acquisition. A general retrieval with the key words "errors in word classes" and "errors in parts of speech" in the CNKI only generated 5 results. But all of them are about Chinese word classes for foreign learners. Simple as they seem to be, EFL learners experience real difficulty in learning and using them. Consequently, they may make errors in word classes in their compositions. Statistics show that the number of errors in word classes accounts for 4.18% of the total tagged errors, and rate the eighth in the 63 types of errors in the Chinese Learner English Corpus (CLEC), which is even greater than that of errors in collocations (Gui Shichun, 2005:13). Therefore, the present study aims to find out Chinese EFL learners' common errors in word classes in their written texts based on the CLEC in an attempt to help shed light on the main problems and difficulties they encounter in the acquisition of the word class.

An Overview of the Errors in Word Classes Made By Chinese College Students

Built by Gui Shichun in 2002, the CLEC was a tagged learner corpus with about 1 million tokens. It was made up of written texts by Chinese English learners of 5 groups: high school students (st2), lower grade college students (non-English major) (st3), higher grade college students (non-English major) (st4), lower grade college students (English major) (st5) and higher grade college students (English major) (st6). As this

¹Guangdong University of Foreign Studies, Guangzhou, Guangdong, 510420

² The research is sponsored by the 2013 special fund for the disciplinary development at institutions of higher learning in Guangdong Province (No. 2013WYXM0037).

study focuses on the errors made by college students, only the material from st3, st4, st5 and st6 were collected.

Errors in word classes are tagged as [wd2] in the CLEC. Through data retrieval, we got 2,596 [wd2] tags in the corpus, and the number of errors made by each group of st3, st4, st5 and st6 is 1,081, 937, 269 and 309 respectively which accounts for 42%, 36%, 10% and 12% as shown in Fig. 1.

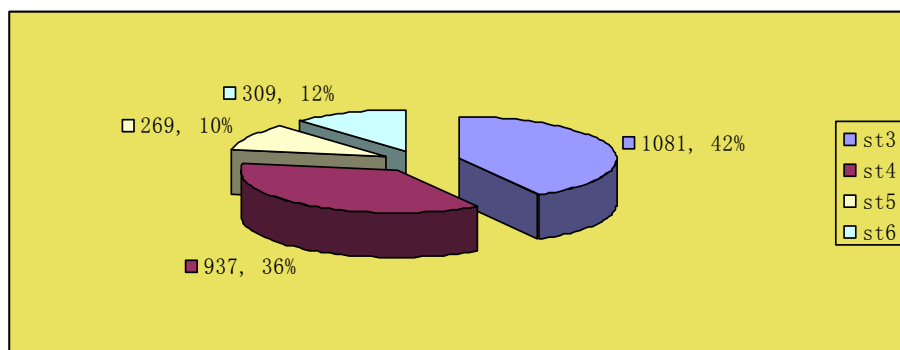


Fig 1: Rate of the Errors in Word Classes for Each of the Four Groups

From Fig. 1, we can see that the error rate in word classes relates closely to learners' English proficiency. Non-English majors (st3 and st4) committed much more errors than English majors (st5 and st6) did (78%:22%); and the lower grade college students (st3) made more errors than higher grade college students (st4) (42%:36%). The less advanced learners (st3) committed more errors compared with the more advanced learners (st4, st5 and st6). However, word classes pose problems for each group of students.

In all the errors in word classes, 688 of them appeared only once which makes up 26.5% of all. 180 of them appeared twice, 73 of them three times and 48 of them four times. This shows that not all the Chinese learners make the same form of errors in the word class.

However, the greatest frequency was 81 for the word *economic*. Next came the word *increasing* (78), *social* (66), *success* (47), *haste* (38), and *developing* (37). So learners made these errors repeatedly. This deserves our attention.

2. Classification of the Errors in the Word Class

In order to have a clearer picture of the errors, we sorted out the first 100 errors with the greatest frequency from all the errors. They are listed in Table 1. The frequency ranges from 4 to 81. It is interesting to find out that the number 11 is a critical point. The number above it has a frequency of more than 20. The frequency decreases markedly for the first 10 errors, but gradually for the rest of them.

Table 1: The first 100 Errors with the Greatest Frequency

No.	Freq.	Word	No.	Freq.	Word	No.	Freq.	Word	No.	Freq.	Word
1	81	economic	26	10	great	51	7	lost	76	5	Chinese
2	78	increasing	27	10	important	52	7	means	77	5	basic
3	66	social	28	10	recently	53	7	health	78	5	past
4	47	success	29	10	limit	54	7	sale	79	5	simply
5	38	haste	30	10	prepare	55	7	adventure	80	5	easily
6	37	developing	31	9	quick	56	7	suit	81	5	hurry
7	34	healthy	32	9	hungry	57	6	wrong	82	5	danger
8	31	live	33	9	global	58	6	perfect	83	5	practice
9	22	raining	34	9	pleasure	59	6	easy	84	5	loss
10	21	studying	35	9	failure	60	6	possible	85	5	graduation

11	18	good	36	9	develop	61	6	industrial	86	5	they
12	17	using	37	8	serious	62	6	guilty	87	5	use
13	17	medical	38	8	rapid	63	6	rich	88	5	comfort
14	16	practice	39	8	careful	64	6	seriously	89	5	understand
15	15	choice	40	8	active	65	6	badly	90	5	require
16	14	society	41	8	really	66	6	firstly	91	5	Practice
17	14	importance	42	8	obviously	67	6	service	92	5	improve
18	14	product	43	8	economics	68	6	image	93	5	fail
19	13	short	44	8	sick	69	6	crime	94	4	successful
20	13	serve	45	8	shortage	70	6	produce	95	4	slow
21	12	difficult	46	8	serves	71	5	Last	96	4	technical
22	12	well	47	8	living	72	5	serving	97	4	stable
23	11	mainly	48	7	patient	73	5	ill	98	4	wealth
24	11	succeed	49	7	interesting	74	5	harmful	99	4	socialism
25	11	mean	50	7	happy	75	5	different	100	4	teach

We further classified these 100 errors according to the misused word class of these words. In Ex. 1, the adjective *broad* was wrongly used as a verb in the sentence; and the right word should be broaden. This kind of error will be classified into the group of *adj-v*, meaning an adjective is mistakenly used as a verb. The classification is shown in Table 2.

Ex. 1 studying abroad helps to broad [wd2,2-2]our mind and train our character.

Table 2: The Classification of the First 100 Errors

No.	Word	Type	No.	Word	Type	No.	Word	Type	No.	Word	Type
1	Good	<i>adj-adv</i>	26	global	<i>adj-n</i>	51	badly	<i>adv-adj</i>	76	graduation	<i>n-v</i>
2	Great	<i>adj-adv</i>	27	active	<i>adj-n</i>	52	simply	<i>adv-adj</i>	77	limit	<i>v-adj</i>
3	Quick	<i>adj-adv</i>	28	patient	<i>adj-n</i>	53	easily	<i>adv-adj</i>	78	suit	<i>v-adj</i>
4	Serious	<i>adj-adv</i>	29	interesting	<i>adj-n</i>	54	haste	<i>n-adj</i>	79	use	<i>v-adj</i>
5	Rapid	<i>adj-adv</i>	30	happy	<i>adj-n</i>	55	society	<i>n-adj</i>	80	comfort	<i>v-adj</i>
6	Careful	<i>adj-adv</i>	31	possible	<i>adj-n</i>	56	importance	<i>n-adj</i>	81	live	<i>v-n</i>
7	Wrong	<i>adj-adv</i>	32	industrial	<i>adj-n</i>	57	pleasure	<i>n-adj</i>	82	practise	<i>v-n</i>
8	Perfect	<i>adj-adv</i>	33	guilty	<i>adj-n</i>	58	sick	<i>n-adj</i>	83	serve	<i>v-n</i>
9	Easy	<i>adj-adv</i>	34	serving	<i>adj-n</i>	59	shortage	<i>n-adj</i>	84	succeed	<i>v-n</i>
10	Last	<i>adj-adv</i>	35	ill	<i>adj-n</i>	60	health	<i>n-adj</i>	85	prepare	<i>v-n</i>
11	successful	<i>adj-adv</i>	36	harmful	<i>adj-n</i>	61	hurry	<i>n-adj</i>	86	develop	<i>v-n</i>
12	Slow	<i>adj-adv</i>	37	different	<i>adj-n</i>	62	danger	<i>n-adj</i>	87	serves	<i>v-n</i>
13	economic	<i>adj-n</i>	38	Chinese	<i>adj-n</i>	63	wealth	<i>n-adj</i>	88	produce	<i>v-n</i>
14	increasing	<i>adj-n</i>	39	basic	<i>adj-n</i>	64	socialism	<i>n-adj</i>	89	understand	<i>v-n</i>
15	Social	<i>adj-n</i>	40	technical	<i>adj-n</i>	65	success	<i>n-v</i>	90	require	<i>v-n</i>
16	developing	<i>adj-n</i>	41	lost	<i>adj-v</i>	66	choice	<i>n-v</i>	91	Practise	<i>v-n</i>
17	Healthy	<i>adj-n</i>	42	rich	<i>adj-v</i>	67	product	<i>n-v</i>	92	improve	<i>v-n</i>
18	Raining	<i>adj-n</i>	43	past	<i>adj-v</i>	68	failure	<i>n-v</i>	93	fail	<i>v-n</i>
19	Studying	<i>adj-n</i>	44	stable	<i>adj-v</i>	69	sale	<i>n-v</i>	94	teach	<i>v-n</i>
20	Using	<i>adj-n</i>	45	well	<i>adv-adj</i>	70	adventure	<i>n-v</i>	95	mean	<i>v-n</i>
21	Medical	<i>adj-n</i>	46	mainly	<i>adv-adj</i>	71	service	<i>n-v</i>	96	living	life
22	Short	<i>adj-n</i>	47	recently	<i>adv-adj</i>	72	image	<i>n-v</i>	97	firstly	first
23	Difficult	<i>adj-n</i>	48	really	<i>adv-adj</i>	73	crime	<i>n-v</i>	98	they	their
24	important	<i>adj-n</i>	49	obviously	<i>adv-adj</i>	74	practice	<i>n-v</i>	99	means	meaning
25	Hungry	<i>adj-n</i>	50	seriously	<i>adv-adj</i>	75	loss	<i>n-v</i>	100	economics	economy

English word classes are comprised of open classes (nouns, adjectives, verbs, and adverbs) and closed classes (pronouns, prepositions, conjunctions, articles, interjections, and others). From Table 2, we can see that all the errors in word classes are open classes with 44 misused adjectives, 23 nouns, 19 verbs and 9

adverbs. But it doesn't mean closed classes will not be misused by Chinese learners as we only sorted out the first 100 word forms with the greatest frequency.

For adjectives, Chinese learners may misuse them as nouns, adverbs or verbs. Below are the examples extracted from the CLEC with the errors highlighted by boldface.

Ex. 2 so the students must do as good [wd2,-] as they can to reduce his

Ex. 3 en be prone to work in the social [wd2,1-1] and the women stay at

Ex. 4 Because it may make people lost [wd2,-] their lives. Of course we all

Ex. 2 is a case of adj-adv, i.e. the adjective was misused as an adverb. It is very common in Chinese learners' compositions, especially in the context of as...as where the adverb was separated from the verb it modified. In Ex. 3, the noun society should be used instead of the adjective social. It belongs to the type of adj-n. It is the most common type of errors in the CLEC. Ex. 4 is a less common type of error where the adjective was misused as a verb, i.e. adj-v.

Nouns may be wrongly used as adjectives or verbs. Here are the examples from the CLEC:

Ex. 5 to get it too fast, you will failure [wd2,1-]. We said [vp6,-s], haste

Ex. 6 fresh water is relatively shortage [wd2,2-] . Nowadays, the shortage

In Ex. 5, the learner misused the word failure as a verb. He or she intended to say that we sometimes need to slow down, or we will fail. It is a type of n-v error. Ex. 6 is a type of n-adj error where a noun was incorrectly used as an adjective. This is an uncommon case because the noun shortage is much more difficult than the adjective short.

For verbs, learners may misuse them as nouns or adjectives. The v-n type of errors is more common than the v-adj type. Below are the examples from the CLEC.

Ex. 7 it [wd5,1-] will lead to succeed [wd2,-] easily. But another people

Ex. 8 the women's power is limit [wd2,1-] in contrast with the men's pow

In Ex. 7, the verb succeed was used where a noun success should be used instead. In Ex. 8, the learner confused the adjective limited with the verb limit.

The last type of error is adv-adj. In that case, an adverb is inappropriately used as an adjective.

Ex. 9 newspaper is the most simply [wd2,2-] and efficient way. We can

In Ex. 9, the learner used the adverb simply where the adjective simple should be used.

Careful readers might have found that the last 5 words listed in Table 2 had not been assigned an error type. These five words are firstly, they, means, living and economics. Below are the sentences taken from the CLEC:

Ex. 10 If I were you, I would make a plan firstly [wd2,-]. Then I would finish it step

Ex. 11 o [wd5, s -] eating! They also build they [wd2,3-1] industries, developing [vp4

Ex. 12 ery and dehumanization, then, the means [wd2, 1-] of euthanasia should be

Ex. 13 e is more distinct. Because the economics [wd2,1-3] of developing countries is

Ex. 14 30 years, people can live a good living [wd2,2-] and have a good medical [w

In Ex. 10, the learner used firstly instead of first, both of which belonged to the same word class. It is the same for the other four words. Therefore, although they were tagged as errors in word classes in the CLEC, they were not rated as an error in the word class in the present study.

To sum up, we have 8 types of errors in word classes. They are adj-adv, adj-n, adj-v, n-v, n-adj, v-n, v-adj, and adv-adj. Among the 8 types of errors, the type adj-n is the most common one. Besides, adjectives can be misused as all the other members of open classes. Therefore, adjectives are most likely to be wrongly

used by Chinese learners. Next are nouns which can be misused as verbs or adjectives. And verbs can also be wrongly used as nouns or adjectives. Finally are adverbs which can be improperly used as adjectives.

3. Possible Causes of the Errors

As stated above, Chinese learners made more errors in word classes than those in collocations. They are very common in their writings. Many factors attribute to the errors.

Difference in the Use of an Adjective between English and Chinese

In general, an adjective is a modifier used to modify a noun or noun phrase. However, an adjective can also be used as a noun or noun phrase in a sentence both in English and Chinese. The difference lies in that in Chinese an adjective can be used as a noun directly, but in English the definite article or other determiners must be added before it. For example, in Chinese one can say “chengshi (honest) shi (is) yizhong (a) meide (virtue)” in which *chengshi* is an adjective meaning *honest*. The correct English version of this sentence is “Honesty is a virtue”. While in English one can say “the old and the young marched together” in which the definite article is used before the adjective to refer to a particular group of people who are old or young.

Another difference in the use of an adjective between English and Chinese is that an adjective in Chinese can be used to perform the function of a disjunct. However, in English, a disjunct is an adverb or adverbial that is loosely connected to a sentence and conveys the speaker's or writer's comment on its content, truth, or manner. For example, one can say “xianran (obvious), daan (answer) shi (is) kendingde (positive).” in Chinese whose English version is “obviously, the answer is positive”. Below are the examples extracted from the CLEC.

Ex. 15 Obvious [wd2,s-] these people compare television with opium.

Ex. 16 General [wd2,-4] speaking, men are taking a leading role in th

Ex. 17 General [wd2,-1] speaking, after the study in [wd3,-2] the sch

Ex. 18 General [wd2, -1] speaking whether would I change my job depen

Ex. 19 But I think, general [wd2, 1-] speaking, if you are good enough, there will

Ex. 20 But general [wd2, 1-] speaking, the more education you receive, th

In these sentences, adjectives were mistakenly used to perform the function of a disjunct where adverbs should be used.

Difference in the Conversion of the Word Class between English and Chinese

English and Chinese differ from each other in the conversion of the word class. In English, it is often accomplished by inflection. For example, the noun form of the verb *punish* is *punishment*. However, in Chinese, a same word form can be used both as a noun or a verb without any change of the word form. For example, in the sentence “ta (he/she) shoudele (be) chengfa (punish)”, *chengfa* is a verb. Its English version is “he/she was punished”. In the phrase “dui (for) fubai (corruption) guanyuan (official) de (of) yanli (tough) chengfa (punishment)”, *chengfa* is a noun. Its English version is “tough punishments for officials convicted of corruption”. Below are the examples extracted from the CLEC.

Ex. 21 he also can think about to choice [wd2,1-] his futur's [fm1,-] creer [fm1,-] .

Ex. 22 There are many ways for us to choice [wd2,1-] , such as the people, the 7

Ex. 23 If I am given a chance to choice [wd2,-] I would prefer the [np9,-] teaching.

Ex. 24 s realities. [sn8,s-] I don't have much chance of choice [wd2,1-] work.

Ex. 25 and this makes it capable [wd3,s-] that we choice [wd2,s-] our jobs.

Ex. 26 People can choice [wd2,s-] one at [wd3,-3] their different characters.

- Ex. 27 One can choice [wd2,-2] any job from hundreds of jobs. But someone li
Ex. 28 Then I will choice [wd2,-s] a fivourist [fm1,-] job to do for a long time
Ex. 29 believe yourself, and choice [wd2,-s] the mode of life you want.
Ex. 30 then choice [wd2,-2] a constant job which I loved [vp6,-s] very mu
Ex. 31 In my opions [fm1,-] , a person should choice [wd2,-] a job in terms of
Ex. 32 There are many kinds of sports that we can choice [wd2,-]. In my
Ex. 33 achuting and sky-ballet may be the sports we can choice [wd2,-] .
Ex. 34 which means not only the employees can choice [wd2, 1-] our graduate
Ex. 35 but also our graduate [np6, s-] can choice [wd2, 1-] employees.

In Chinese, *xuanze* can be used as a noun and a verb. It has two equivalents in English: the noun form *choice* and the verb form *choose*. Ex. 21 to Ex. 35 are all the cases of misusing the noun *choice* as the verb *choose*.

Difference in the Use of a Verb between English and Chinese

In Chinese, a verb can be followed by another verb directly without any conjugation of the latter one. For example, it is grammatical to say “wo (I) feichang (very much) xihuan (like) da (play) lanqiu (basketball)”. But in English, we are required to change the second verb *play* into *playing* or *to play*. So the correct English version of the above Chinese sentence is “I like playing/to play basketball very much”. Besides, in an English sentence, the verb after a preposition must be in the form of a gerund. Below are the examples taken from the CLEC.

- Ex. 36 I want **succeed** [wd2,-] as all the people
Ex. 37 Thirdly, everyone has right to choose **live** [wd2, 1-2] or die, the traditional
Ex. 38 the progress of economy in the **develop** [wd2,1-1] countries led [vp1,-4] the life
Ex. 39 If you want to haste [wd2,2-], the chance of **succeed** [wd2,1-] will go
Ex. 40 we must be very careful, in case of **fail** [wd2,-]

In Ex. 36 and Ex. 37, the verb *succeed* and *live* should be in the form of an infinitive. It is ungrammatical to put a second verb directly after the predicate verb. In Ex. 38, the verb *develop* should be in the form of “developing”. Ex. 39 and Ex. 40 are both the case of using a verb directly after a preposition where a gerund is grammatically required.

Similar Spellings of the Words

Some English words are spelt in a similar way, which will be difficult for learners to distinguish them, especially for less advanced learners. As a result, they may confuse one word with another in their writings. If the wrongly used words belong to different word classes, then they were classified as errors in word classes in the CLEC.

- Ex. 41 Maybe in American people **opposite** [wd2, 1-] euthanasia not because of
Ex. 42 tell the children that there is a **fairly** [wd2, 1-] lady dressed in white
Ex. 43 hearing his hard deep **breathe** [wd2, 1-] sound, I was suffering with him.
Ex. 44 but just give me some **advise** [wd2, 1-] I have right to keep my friend as my
Ex. 45 writing to you to make a **complain** [wd2, 1-] on the noise that disturbs me.
Ex. 46 Because we are in **hasty** [wd2, 1-], we have a press [wd3, 2-], we can't do

In Ex. 41, the learner wrongly used *opposite* as *oppose*. These two words are spelt similarly, but they don't relate to each other in their meaning. The same is true of Ex. 42 in which the mistakenly used word

fairly has nothing to do with the word *fairy* in their meaning. However, from Ex. 43 to Ex. 46, the learners misused the words which are same in their meaning. Actually, they belong to the same word family, for instance *breath* and *breathe*, *advice* and *advise*.

Transfer of Training

Although the word class is the basic element of English grammar, it is overlooked by English teachers in their teaching. They take it for granted that English learners will naturally acquire word classes without any particular efforts. Other grammatical knowledge is much more important than the word class. So they don't give their students enough training in the word class. Moreover, the teaching material also overlooks the word class. They don't provide learners with sufficient exercises in word classes. Let's take the Chinese-English dictionary as an example. This kind of dictionary is mainly used for production. However, the English equivalents in these dictionaries are not provided with a word class label. As a result, the learner may misuse them in word classes.

In sum, Chinese learners may commit errors in word classes for various reasons. From what is stated in section 4, we can see that interlingual transfer might account for most of the errors. As there are substantial differences between Chinese and English in regard to the word class, Chinese learners may inappropriately apply their mother tongue rules and knowledge to their English learning. Consequently, they make errors in word classes. Apart from mother tongue interference, intralingual transfer may also cause Chinese learners to make errors in word classes. When they learn a foreign language, they tend to apply what they have learned to a new item. If they overgeneralize certain target language rules, errors arise.

4. Conclusions

From what is stated above, we may reach the following conclusions. First of all, Chinese learners tend to commit errors in the word class systematically. Generally speaking, less advanced learners make more errors than more advanced learners do. But all the four groups of college students commit errors in word classes. So these errors are not sporadic but repeated in a systematic way.

Secondly, errors in word classes are very common among Chinese English learners. Most of the errors are the words of the open classes. The most common type of error falls into the category of *adj-n*. That is to say, Chinese learners tend to misuse an adjective as a noun. In other words, the adjective poses the greatest problem for Chinese learners, not the noun or verb as we thought it to be.

Thirdly, the acquisition of the word class has long been overlooked in English teaching and learning. Both teachers and learners haven't paid enough attention to it. Therefore, we should reassess the problems that word classes present Chinese learners with, and give special attention to it.

Last but not the least, possible measures can be taken to help learners acquire the word class. Due to the limited space, we'll focus on improvements in teaching material. For example, all the inflected forms of the same word family can be listed under the main entry of the new words in the textbooks. This will not only help learners learn the new word, but also enlarge their vocabulary. In addition, more exercises in the word class can be added in the textbooks. Finally, as an important aid to language learning, the dictionary should cater for the user's needs. As the Chinese-English dictionary is especially made for active use, the dictionary should provide the equivalents with the word class labels, especially when the head word has more than one word class. This will reduce the errors in the word class.

Acknowledgements ; The research is sponsored by the 2013 special fund for the disciplinary development at institutions of higher learning in Guangdong Province (No. 2013WYXM0037).

References

Corder, S. P. 1981. *Error analysis and interlanguage*[M]. Oxford: Oxford University Press.

- Crystal, D. 2003. *The Cambridge Encyclopedia of the English Language* (2nd ed.) [M]. Cambridge: Cambridge University Press.
- Ellis, R. 2008. *The Study of Second Language Acquisition* (2nd edition) [M]. Oxford: Oxford University Press.
- Gui, Shichun, Yang, Huizhong. 2002. *Chinese Learner English Corpus*. Shanghai: Shanghai Foreign Language Education Press.
- Gui, Shichun. 2005. An analysis of speech errors of Chinese learner[A]. In Yang Huizhong, Gui Shichun, Yang Dafu (eds.) *An Analysis of Chinese Learner's Interlanguage based on CLEC*[C] (pp.1-109). Shanghai: Shanghai Foreign Language Education Press.
- Kolln, M. J. & Funk, R. W. 2008. *Understanding English Grammar* (8th ed.) [M], Longman.
- Selinker, L. 1992. *Rediscovering Interlanguage*[M]. London & New York: Longman.
- Selinker, L. & Lakshmanan, U. 1992. Language transfer and fossilization: the multiple effects principle[A]. In Gass, S. M. & Selinker, L. (eds.) *Language Transfer in Language Learning* [C]. Amsterdam/Philadelphia: John Benjamins, 197-216.
- Tao Xiaoyue. *A Corpus-based Typological Analysis of the Written Errors Committed by Chinese English Majors* [D]. Yangzhou University, 2006.