A Corpus-based Study on Chinese EFL Learners' Acquisition of English Existential Construction[*]

Xiaorong Zhang

Northwest Agriculture & Forestry University, Yangling, China

Abstract—**This corpus-based study examines English existential construction used by intermediate and advanced level Chinese EFL (English as a Foreign Language) learners. The corpora adopted are the sub-corpora of SWECCL2.0—TEM 4 Oral and TEM 8 Oral and a sub-corpora of COCA. The study concludes Chinese EFL learners tend to overuse English existential construction and prefer the basic tenses, simple intransitive verbs and commonly seen expressions, avoid the perfect tenses and the difficult forms. With the level of proficiency in English getting higher, Chinese students try to use fewer English existential sentences. The study also finds errors relevant to the tenses, agreements and misuse of "there + have" pattern made in TEM 4 Oral and more difficult participle errors made in TEM 8 Oral and the main reason of errors made in English existential acquisition is due to L1 transfer/ L1 influence**

Index Terms—**English existential construction, acquisition, error analysis, corpus-based, L1 transfer**

I. INTRODUCTION

Existential sentence as a special syntax and semantic structure has been widely used by Chinese EFL learners. However, various errors occurred in Chinese EFL learners' acquisition of English existential sentences.

In Sasaki's (1990) study of 173 Japanese learners' composition, she found that besides the more native-like *there be* construction which was approximately 30%. Sasaki also noted that there was a general shift from topic-comment to subject-predicate structure as students' proficiency increases.

Yip (1995) investigates some problems in Chinese learners' use of existential. However, her subjects are exchange students from Hong Kong and Singapore. These students could not really represent Chinese EFL learners as a whole. What's more, her study was confined to a small amount of data and a qualitative study.

In order to study English native-speakers' and the Chinese EFL learners' usage of EEC (English Existential Construction), Song Jingsheng (2006) compared the collected data from BNC and CLEC corpus. By analyzing and comparing distributional frequencies of different types of EEC in these two corpora, he found there are regularity and some characteristics in their distributional frequencies. In the process of negation usage, there is a significant difference between native speakers and Chinese learners. But Song Jingsheng's research just used frequency counts to answer his study questions.

Dai Manchun and Liang Yi (2007) analyzed the syntactic features of EEC from minimalist program theory and used empirical approach to study Chinese students' acquisition of EEC. They divided the EEC into five types and targeted on four group participants who included high school and university students. In the research, Dai stated there were different results of Chinese students' acquisition of different EEC because of EEC's syntactic operation. That is to say, for the more simple syntactic operation of EEC, the higher degree of mastery.

From the discussion above, researchers have made lots of contributions to the study of existential sentences in theory and in SLA, which provide useful views for the present study. These studies either focused on a stage of proficiency, or just did simple frequency counts. Related studies from corpus approach are still inadequate in exploring Chinese EFL learners' acquisition of English existential Construction.

II. RESEARCH DESIGN AND METHODOLOGY

A. Research Questions

1. In comparison with native English speakers, what features are presented in Chinese EFL learners' acquisition of English existential construction?
2. What is the pattern of Chinese EFL learners' errors of existential construction at different levels?
3. What are the causes of the errors made by Chinese EFL learners?

B. Corpus and Tools Used in the Study

The data is collected from two corpora. One represents Chinese EFL learners' speaking and the other is

native-speaker corpus as reference. The former is SWECCL (Spoken and Written English Corpus of Chinese Learners) and the latter is COCA (Corpus of Contemporary American English). The tool adopted in the study is AntConc 3.2.1, which is green and cross-platform corpus processing software developed by Laurence Anthony. The software has a variety of functions including Concordance, Concordance Plot, File View, Clusters, Collocates, Word List and Key Word List.

C.   Procedures for Data Processing

Step 1. See "there" as the key word, use AntConc3.2.1 to search all the sentences containing "there" in SWECCL which include TEM4 Oral and TEM-8 Oral and part of sample of spoken COCA is dealt with online in the same way.

Step 2. Non-existential sentences and vaguely expressed ones should be excluded and the frequency of existential sentences in the three corpora should be counted one by one. Owing to the difference in size of the three corpora, the original frequency should be standardized to make the comparison more acceptable. The formula of frequency standardization of per million tokens is (original frequency/corpus tokens) x1000,000 (Biber et al 2000).

Step 3. Classify English existential construction into five categories: "there + be" "there + modal be", "there + marginal modal be" "there + semi-auxiliary be" and "there + intransitive verb". Divide five types into two main types based on simple or complex grammatical structure.

Step 4. Compare the frequency of all existential sentences and of different forms between TEM 4 Oral, TEM 8 Oral and part of sample of SPOKEN in COCA to find out the characteristics of Chinese EFL learners' acquisition of English existential sentences.

Step 5. Collect existential sentences with errors in TEM 4 Oral and TEM 8 Oral. Next, ignore the obviously meaningless spelling errors, including those omitting the predicate verb "be" and some not belonging to errors of the properties of existential sentences such as NP (noun phrase), PP (preposition phrase). Then calculate all errors of there-existential structure in both corpora.

Step 6. Classify the errors into agreement, tense, finite/non-finite, structural deficiency. Use Antconc3.2.1 to search the frequency of these types of errors in TEM 4 Oral and TEM 8 Oral. And then calculate them.

Step 7. Compare the frequency of different types of errors between TEM 4 Oral and TEM 8 Oral to explore the regularity of Chinese EFL learners' acquisition of English existential construction. Causes for the errors are analyzed to help explain sources of the regularity and improve learning strategy and learners' awareness.

Step 8. From the analyses of the characteristics and regularity of Chinese EFL learners' acquisition of English existential construction, explore the solution to avoid errors made in the acquisition process.

III.   DATA ANALYSIS AND DISCUSSION

A.   Characteristics of Chinese EFL Learners Acquisition of Existential Construction

1. The Difference in the Frequency of Existential Construction between Chinese EFL Learners and Native English Speakers

TABLE 1
FREQUENCY AND STANDARDIZATION OF ENGLISH EXISTENTIAL CONSTRUCTION IN TEM 4 ORAL,
TEM 8 ORAL AND PART OF SAMPLE OF SPOKEN IN COCA.

	TEM 4 Oral	TEM 8 Oral	COCA
Tokens	779731	286583	554526
Frequency	3105	1013	1686
Standardization	3982.14	3534.75	3040.43

(The formula of frequency standardization of per million tokens is (original frequency/corpus tokens) x 1000,000 (Biber et al 2000).)

From table 1 it can be seen the number of existential sentences in TEM 4 Oral or TEM 8 Oral is much higher than that of COCA, which means Chinese EFL learners may tend to overuse existential sentence. It can also be observed that the number of existential sentences in TEM 4 Oral is higher than in TEM 8 Oral, which shows the higher level of Chinese EFL learners are, the fewer existential sentences they use in speaking.

2. The Difference in the Use of Various Forms of Existential Construction between Chinese EFL Learners and Native English Speakers

According to the five types of existential construction in the front, it is divided into the basic forms and other forms. The basic form are "there + be" and "there + modal be" patterns. Other forms are "there + intransitive verb" "there +semi-auxiliary be" and "there + marginal modal be" patterns.

TABLE 2
FREQUENCY OF THE BASIC FORMS OF ENGLISH EXISTENTIAL CONSTRUCTION IN TEM 4 ORAL, TEM 8 ORAL AND COCA

Basic forms of Existential Sentences			TEM 4 Oral	TEM 8 Oral	COCA
there + be	present tense	there is	660	192	185
		there are	552	465	195
	past tense	there was	467	19	128
		there were	108	20	53
	perfect tense	there has been	0	2	8
		there have been	2	0	14
		there had been	0	1	3
There+ modal be		there will/ would/ can/could/may/might/should/must be	174	96	68
TOTAL			3961	795	654

From Table 2, it can be found that the frequency of the basic form in TEM 4 Oral and TEM 8 Oral is much higher than that in COCA. For the basic form of " there be " structure, Chinese EFL learners tend to use the most of present tense and past tense and the least in perfect tense.

For the "there + modal be" form, "will" "can" "must" are the major choice of the TEM 4 Oral. "Will" "should" "must" are the major choice of the TEM 8 Oral. Of all the modal verbs in TEM 4 Oral and TEM 8 Oral, "will" is used 140 times in TEM 4 Oral and 66 times in TEM 8 Oral. However, in COCA, modal verbs are used in balance.

TABLE 3
FREQUENCY OF OTHER FORMS OF ENGLISH EXISTENTIAL CONSTRUCTION IN TEM 4 ORAL, TEM 8 ORAL AND COCA

Other forms of English existential construction		TEM 4 Oral	TEM 8 Oral	COCA
there + Intransitive verb	there lives/lived	1		
	there stands/stood			
	there comes/came/will come	7	2	
	there exists/existed	1		1
	there lies			1
	there follows			1
There + semi- auxiliary be	there is going to be	2		2
	there is likely to be			1
	there is found to be			1
	there is expected to be			
	there is said to be			
there + marginal modal be	there seems/ed to be	1	1	1
	there needs to be			1
	there used to be	1		1
	there remains to be			
	there appears to be			
TOTAL		13	3	10

As for the other forms shown in table 3, the frequency of all the structures in TEM 4 Oral, TEM 8 Oral and COCA is kept nearly the same. Amazingly, there are more structures of the other forms in TEM 4 Oral. However, it is easy to find that the structures are peculiar to some commonly seen ones, "there comes/came/will come", "there is going to be". In TEM 8 Oral, there appears another two structures "there comes/came/will come", "there seems/ed to be". Compared with TEM 4 Oral and TEM 8 Oral, flexible use of more predicate verbs and structures enrich the other forms in COCA. For instance, "there exists/existed", "there lies", "there follows", "there is going to be", "there is likely to be", "there is found to be", "there seems/ed to be", "there needs to be", "there used to be".

It is easy to see that the present and past tense used in the basic forms or "live" "come" and "exist" appearing in the pattern "there + intransitive verb" "be going to be" in the "there + semi-auxiliary be", "seem to be" and "used to be" in the "there + marginal modal be", which belong to the other forms, are the language items with high frequency. As mentioned above, language input frequency can influence language output ability to some extent.

B.   The Analysis of Chinese EFL Learners' Acquisition of Existential Construction

Although there exist many differences between TEM 4 Oral and TEM 8 Oral, errors appear in both but with different number and different types. TEM 4 Oral and TEM 8 Oral are not fully error tagged, thus errors in the existential have to be collected semi-manually with the aid of KWIC and File View accompanied by the non-error context. The types of errors are divided into agreement, finite/ non-finite, tense and structural deficiency.

1. The Difference in the Number and types of Errors between TEM 4 Oral and TEM 8 Oral

TABLE 4
FREQUENCY OF DIFFERENT TYPES OF ERRORS IN TEM 4 ORAL AND TEM 8 ORAL

Types of Errors		TEM 4 Oral	TEM 8 Oral
1 tense		40	7
2 agreement		25	9
3 finite/ non-finite		7	10
structural deficiency	4 there have	15	5
	5 pseudo-relatives	11	9
	6 others	12	8
TOTAL		110	93

From table 4, it can be seen that more errors of tense, agreement, and the misuse of the "there have" pattern appear in TEM-4 Oral while there are more errors of finite/ non-finite in TEM-8 Oral. The number of pseudo-relatives and other errors classified into structural deficiency is similar in these two sub-corpora. Standardization of the frequency of different types of the errors in figure 2 directly shows the tendency that types of errors vary in the two sub-corpora.

2. Analysis of Different Types of Errors between TEM 4 Oral and TEM 8 Oral

(1) Misuse of Tense

The appropriate use of various tenses in English is always a headache for Chinese English learners. Misuse of tense in the *existential* clause in TEM 4 Oral and TEM 8 Oral mainly includes the misuse of the present tense for the past tense, the misuse of past tense present tense and the misuse of present tense and past tense for the perfect tense. Examples are selected and presented as follows:

TEM 4 Oral:

Example 1: The doctor told me that there is no hope that my mother will be waken.

Example 2: There is a little cat in the middle of the street and when... he said somebody will hurt the cat.

TEM 8 Oral

Example 3: It told us that there are still er ... very many many er ... issues that we should overcome er ... we should protect our pilots.

English is a language that possesses 16 tenses and each tense is expressed by the corresponding form of the predicate verb. However, in Chinese there is no notion of tense but the concept of time – *xianzai* (present), *guoqu* (past), and *jianglai* (future) which is expressed by simply adding the time adverbial such as *jintian* (present), 2015, *henjiu yiqian* (long time ago) with no transformation in the predicate verb, that is to say, in Chinese the time adverbial and the predicate verb are two totally independent notions and the change of one cannot influence the other. The obvious difference in tense between Chinese and English results in high degree of difficulty for Chinese English majors in learning and mastering English tenses.

To explain the phenomenon that present tense is overly misused for other tenses, we should also resort to induced interference. Due to its simplicity, the present tense is first introduced and exposed to Chinese English learners and is practiced a lot by Chinese English learners. In addition, the present tense is widely used in daily communication, argumentation writing and so on. The induced teaching and the excessive exposure and drilling altogether lead to Chinese English majors' misuse of the present tense for other tenses.

(2) Disagreement of a Subject and Predicate

In English the subject and the predicate verb must agree. In the existential clause, subject-verb agreement reflects differently from two aspects. One is that it is the postponed NP as the logical subject that determines subject-verb agreement, rather than "there" functions as the surface subject. The other is that when "there" is followed by conjoined noun phrase subjects, existential clause obeys the proximity principle, that is, the verb BE agrees with the number of the nearest conjunct rather than the number of both noun phrases combined (Murcia & Larsen-Freeman 1983). The following examples present clearly this principle:

Example 4: There is a student and two teachers in the classroom.

Example 5: There are two teachers and a student in the classroom.

However, in spoken English "there" is perceived as a singular subject, and the logical subject is ignored for subject-verb agreement purposes (Murcia & Larsen-Freeman 1983). Subject-verb disagreement is frequently committed by Chinese English majors. Some examples are extracted from TEM 4 Oral and TEM 8 Oral and presented below.

TEM 4 Oral:

Example 6: But you know there is some examples.

Example 7: You know there was many many air crashes.

Example 8: One night, there were a fire on the house.

TEM 8 Oral:

Example 9: There are someone in it.

One possible explanation for subject-verb disagreement may be interlingual interference. On one hand, the difference between meaning-focused Chinese and form-focused English interferes the mastery of subject-verb agreement in existential. As is known to all, it is the nearest conjunct of the subjects that determines subject-verb agreement when existential possesses conjoined noun phrase subjects. However, Chinese EFL learners who are deeply influenced by Chinese naturally and unconsciously consider conjoined noun phrase subjects as the determiner of subject-verb

agreement.

(3) Finite/Non-finite Errors

The major problem in TEM 8 Oral is finite/non-finite and the structure with the error is "There be +somebody/something + v-ing/v-ed + locative adverbial". As is known, although the forms of the present participle are usually regular, the forms of the past participle can be regular or irregular. Besides, there are various irregularities. The general grammatical rule to use present and past participle is clear that the present participle is used to express the active voice while the past participle is linked to the passive voice. This type of errors in TEM 8 Oral shows as proficiency of the target language increases, the learners are willing to try some complex structures. Few errors of finite/non-finite in TEM 4 Oral do not mean that students are good at using the grammar. Instead, they may employ avoidance strategy to deal with the difficulty in learning the participles, so actually there are only a very small number of existential sentences.

(4) Existential Pseudo – relatives

Sentences taking on the form of "there be +NP+VP" are called *existential pseudo-relatives* by Yip (1995). It is observed that this kind of sentence lacks relative pronoun. Some examples of *existential pseudo-relatives* from TEM 4 Oral and TEM 8 Oral are presented as follows:

TEM 4 Oral:

Example 10: And because there are many people stay in the train station.

TEM 8 Oral:

Example 11: Now there are many students study in the classroom.

Yip (1995) laid stress on the overuse of *existential pseudo-relatives,* for example, "There are many people study English" which, she thought, was caused by Chinese learners' directly copying the pivotal constructions of Chinese existential sentences. But English existential structure owns the only predicate verb "be" without any other verbs following unless they are in relative clauses or appropriate participle forms, thus this Chinese existential sentence should be expressed as " There' re many people who study English" or "There are many students studying English".

(5) Misuse of the "There have" Pattern

In regard to the question why many students in TEM 4 use "there + have" instead of "there + be", the reason may lie in that they confuse the "there + be" pattern with the content verb "have" both of which can express the meaning of the Chinese word "*you*" (have). The Chinese word "*you*" (have) is also translated into the English verb "have". It is Chinese thinking pattern that gives rise to this type of errors. At the same time, students doesn't grasp the meaning of "there be" pattern but use "have" which means "own" to express the meaning of existence. It appears the transfer of the vocabulary thus expand the usage of "have" and reduce the function of "there be". These errors are affected by the typical Chinese pivotal construction. Chinese students at intermediate level and high level are both inclined to follow it. As a result, this kind of error cannot be easily removed and has developed into a fossilized language phenomenon.

IV. CONCLUSION

A. Findings

Firstly, compared with native speakers, there is tendency that Chinese EFL learners overuse English existential construction and the distribution of different forms of the structure is rather unbalanced. In addition, Chinese EFL learners like using the basic form of the structure including "there + be" and "there + modal be". What's more, the predicate verb "be" is mainly in the present tense and past tense, but the perfect tense is even rarely used. For the other forms, Chinese EFL students also concentrate on some commonly used words and expressions. For instance, in TEM 4 Oral and TEM 8 Oral the predicate verbs of the structure "there + intransitive" are limited to the simple words "comes/came/will come". "There is/was going to be" is the only choice for the structure "there +semi-auxiliary be" in TEM 4 Oral. "There seems to be" and "there used to be", the representatives of the structure "there + marginal modal be", are only found in TEM 4 Oral. Therefore, the acquisition of existential construction mainly focuses on basic elements and lacks flexibility.

Secondly, through comparing and analyzing the errors of English existential construction in TEM 4 Oral and TEM 8 Oral, it is found errors in TEM 4 Oral are more than that of TEM 8 Oral and the errors in both sub-corpora cover nearly all the types relevant to the properties of the structure, such as the disagreement of subject and predicate and the confusion of "there + be" with "there + have". The major errors of TEM 4 Oral is the basic grammatical rules such as tense, agreement and the "there + have" pattern while the finite/ non-finite which is the more complex structure is the major error in TEM 8 Oral. They also have the common errors in pseudo-relatives and other structural errors. From various errors, it shows Chinese EFL learners' acquisition of English existential construction conforms to the general regularity of human cognition, developing in order from simplicity to complexity, from elementary to advanced.

Lastly, several factors are found to account for these errors, the main factors are: L1 transfer, over generalization and developmental factors. Generally speaking, learners with low proficiency level generate more *there*-sentences than learners with higher proficiency. There are two main factors that influence Chinese learners' over generation of *there*-sentences. One is the influence of L1. Chinese is topic-prominent language while English is subject-prominent language. This phenomenon is caused by transfer from learner's first language. Chinese learners are especially influenced by Chinese *you*-sentence. The other factor is developmental factor. With the development of learners'

proficiency level and learners' gradual perception and understanding of the L2 rules, learners employ various sentence structures and produce fewer *there*-sentences. The other one is affected by the typical Chinese pivotal construction. Chinese students at intermediate level and high level are both inclined to follow it. As a result, this kind of error cannot be easily removed and has developed into a fossilized language phenomenon that is even accepted in second language acquisition. For example, "*Jiao Shi Li You Ji Ge Xue Sheng Zai Du Shu*" (Some Students are reading in the classroom), is likely to be translated into "There are some students read in the classroom.".

B.   Pedagogical Implications

For various errors occurred in Chinese EFL learners' acquisition of English existential construction, spoken English should be paid more attention in the future teaching. The teaching of any language points should be systemic, complex *there*-sentence structures with relative clause, participles and infinitives should be taken as target grammar items to present learners. Some teaching and learning strategies need to be adopted to help students understand the correct way of using *there*-structure. Teachers and students should make clear of the frequently used "there be" structure and the ones that do not appear in natives' spoken. In most text-books or grammar books used by Chinese teachers and students, these items are not given a systematic teaching. However, if teachers' instruction should cover complex *there*-sentences, and enable learners to get a full picture of *there*-sentences, they may be able to avoid producing many pseudo-relatives to a great extent.

In the process of language teaching, teachers should pay much attention to the differences between English and Chinese and use the corpus to present the errors and analyze them. As to overused, underused and misused types, teachers should warn students to cautiously use these types when requiring students to do pattern drills. Since oral output might cause Chinese EFL learners to overuse some abbreviation forms, teachers should point the error out at proper time. Besides, English teaching materials should be as authentic as possible on the ground that the unauthentic teaching material has a negative influence on students' language learning and is one of the reasons for overuse, underuse and misuse in the process of language learning.

C.   Limitation

Due to the limit of time and energy, the present study only chooses and analyzes TEM 4 Oral, TEM 8 Oral and part of sub-corpora of COCA. In terms of the subject in this study, only Chinese English majors are included, who only account for a small part of Chinese EFL learners. Therefore, the findings may not be a reflection of the characteristics in other groups of Chinese EFL learners, and the tentative explanations may not be able to apply to problems of other learners. Besides, the study is only corpus-based, future research should include more designed experiments so as to make the finding more objective and convincing.

REFERENCES

[1]   Celce-Murcia, M., & Larsen-Freeman, D. (1983). The Grammar Book. Rowley, MA: Newbury House.
[2]   Chomsky, N. (1981). Lectures on Government and Binding. Dordrecht:Foris.
[3]   Dai Manchun (2000). The Study on English Existential Construction. *Foreign languages and Their Teaching.9, 24-27.*
[4]   Dai Manchun, Liang Yi. (2007). The Acquisition of English Existential Constructions by Chinese EFL Learners. *Foreign Languages Research* 6, 41-48.
[5]   Feng Tao, Shen Gang. (2011). An Empirical Study of the Acquisition of *There*-Construction by Chinese Non-English major EFL learners. *Journal of Minorities Teachers College of Qinghai Teachers University* 22.1, 62-65.
[6]   Frezz, R. (1992). Existential and Other Locatives. *Language* 68. 3 ,553- 595.
[7]   Gao Wencheng, Zhang Lifang. (2010). Three Issues in Cognitive Contrastive Study of English and Chinese Existential Construction. *Journal of Xi'an International Studies University.* 18(1):23-28
[8]   Gu Yang. (1997). The Existential Sentences in Chinese. *Modern Foreign Languages* 77.3, 14-25.
[9]   Gu Hualing. (1998). The Study of Thematic Structure in English Existential Construction. *Journal of Foreign Languages.*6, 33-38.
[10] Haegeman, L. and Gueron, J. (1999). English Grammar: A Generative Perspective. Oxford: Blackwell.
[11] Halliday, M.A.K. (.2000).An Introduction to Functional Grammar. Beijing: Foreign Language Teaching and Research Press.
[12] Han Jingquan. (2001). Existential Sentences in English and Chinese: towards a Generative Analysis. *Modern Foreign Languages* 2, 143-158.
[13] Huang Bing, He Anping. (2007). The Acquisition of *There-be* by Chinese EFL College students. *Modern Foreign Languages* 30.1, 57-64.
[14] Jin Jiling. (1996). Contrastive Study of Existential Structure in English and Chinese.(1996). *Journal of Foreign Languages.*6, 10-16.
[15] Li Jinglian, Wang Kefei. (2005). The Syntax of Existential Sentences in English and Chinese. *Modern Foreign Languages* 28.4, 350-361.
[16] Quirk, R.et al. (1972). A Grammar of Contemporary English. London: Longman Group Limited.
[17] Quirk, R.et al. (1985). A Comprehensive Grammar of the English Language. London: Longman.
[18] Sasaki, M. (1990). Topic prominence in Japanese EFL students' existential Constructions. *Language Learning* 40(3): 337-368.
[19] Song Jingsheng. (2006). A Corpus-based Study on English Existential-*there* Sentences. *Journal of Jiangsu University of Science and Technology (Social Science Edition)* 6, 85-89.
[20] Ward, G. & Birner B. (1995). Definiteness and the English Existentials. Language 71, 4, 722-742.

[21] Wen Qiufang, Wang Lifei, Liang Maocheng. (2008). Spoken and Written English Corpus of Chinese Learners. Beijing: Foreign
      Language Teaching and Research Press.
[22] Yip, V. (1995) Interlanguage and Learnability: From Chinese to English. Amsterdam: John Benjamins Publishing Company.
[23] Zhang Zhenbang. (1983). New English Grammar. Shanghai: Shanghai Translation Publishing House.

Xiaorong Zhang is an associate professor at Department of Foreign Languages, Northwest Agriculture & Forestry University.
She received her MA degree in linguistics from Chinese University of Hongkong. Her research interests include second language
acquisition, syntax and studies on public speaking. She is a member of Asia TEFL Association.