An Analysis of Verb Phrase Errors in Chinese Learner English Corpus

Huaqing He

School of Foreign Languages, China West Normal University
Nanchong City, Sichuan Province, China
masha4567@sina.com

Abstract: *This paper employs qualitative and quantitative methods to study verb phrase errors made by Chinese non-English majors in their writing, collected in Chinese Learner English Corpus (CLEC), on the basis of contrastive analysis and error analysis. The purpose is to offer English learners methods to help improve their English proficiency and put forward some suggestions on English language teaching and learning. The main findings are as follows: (1) the verb phrase errors account for 12.08% of the total language errors; (2) there is a significant difference in verb phrase errors made by college learners of different writing proficiency; 3) Three main subtypes of verb phrase errors are errors in agreement, errors in transitivity, and errors in voice. Finally, this paper explores the reasons for verb phrase errors and comes up with some pedagogical suggestions.*

Keywords: *Chinese Learner English Corpus; Chinese English learner; error analysis; verb phrase errors.*

1. INTRODUCTION

During the process of acquiring a native language, learners potentially make different kinds of errors. When learning a foreign language, it is natural and unavoidable for learners to make errors. "If to err and to speak are each uniquely human, then to err at speaking, or to commit language errors, must mark the very pinnacle of human uniqueness" [1]. Corder (1967) stated that errors could be significant in three ways: (1) they provided the teacher with information about how much the learner had learnt, (2) they provided the researcher with evidence of how language was learnt, and (3) they served as devices by which the learner discovered the rules of the target language [2]. Therefore, it is common in foreign language teaching and learning to emphasize the role of errors.

2. LITERATURE REVIEW

With the popularization of computers, the approach of corpus-based research on interlanguage (IL) and language errors is becoming widespread. Professor Gui Shichun and other scholars at Guangdong University of Foreign Studies began to establish a one-million word corpus (Chinese Learner English Corpus) in 1997. They try to conduct systematic tagging and analysis of the learner errors and making a contrastive survey of interlanguage features in various dimensions. The data of this IL corpus is collected from the written production of general subject-matter by Chinese learners from middle schools and high schools up to English majors in university. In this corpus, variables including school, age, sex, level, time of learning, and way of performance are followed in the process of data collection and are annotated to improve comparative analysis of various purposes. In addition, an error tagging system is designed since it is part of the important preliminary work [3].

A lot of researchers have studied language errors of EFL learners' writing In China. Chen Wanxia made a tentative analysis on collocation errors in students' essays written by some junior students at a Foreign Language University in He Nan Province in China. She reached a conclusion that out of the collocation errors, grammatical collocation errors accounted for 55.83%, and lexical collocation errors accounted for 44.17% [4]. Li Jingquan and Cai Jinting used CLEC to study the misuse of English articles and they found that the misuse of articles by Chinese learners mainly followed three patterns: omission of articles, over-application of articles and confusion of articles [5]. Su Hongxia analyzed errors of three high frequency verb phrases "buy, wait and learn" in CLEC [6]. He Huaqing found that Lexical errors in a sub-corpus of CLEC comprised 30% of all written mistakes made by Chinese college students in English and there was a negative correlation between lexical errors and writing quality [7]. Xia discussed possible causes of word class errors in the CLEC [8].

In summary, Chinese researchers conducted a lot of studies on language errors in English writing and made great contributions to English teaching and learning. This paper employs qualitative and quantitative methods to explore verb phrase errors made by Chinese non-English majors in their writing, collected in CLEC, which has the same title "Health Gains in Developing Countries". The purpose is to offer English learners methods to help improve their English proficiency and put forward some suggestions on English language teaching and learning.

3. RESEARCH DESIGN

This study employs contrastive and error analyses to explore verb phrase errors collected from written samples of non-English majors in the CLEC.

3.1. Research Questions

The purpose of this study is to investigate the Chinese non-English majors' verb phrase errors in their writing. Specifically, the present study attempts to answer the following questions:

What is the percentage of verb phrase errors out of all the language errors in the writing sample?

Are there any differences in verb phrase errors between the good writing sample and the poor writing sample?

What are main subtypes of verb phrase errors?

3.2. Research Instruments and Procedures

This study employs the following instruments and Procedures:

3.2.1. Chinese Learner English Corpus (CLEC)

CLEC covers five sub-corpuses, made up of written texts by Chinese English learners of 5 groups: high school students, lower grade college students (non-English majors), higher grade college students (non-English majors) lower grade college students (English majors) and higher grade college students (English majors), and covers 1,207,879 words. This study chooses lower grade college students and higher grade college students' corpuses as research subjects. Because the two corpuses cover both different and same writing tasks, the same writing task which is entitled "Health Gains in Developing Countries" was chosen to do more exact analysis. There were 290 compositions, including 130 lower grade college students' compositions and 160 higher grade college students' compositions. There are 40320 word tokens in total, excluding all error tags and annotations.

3.2.2. Ant Conc3.2.1

This is a corpus software tool which can be used to search key words. CLEC is a tagged corpus which tags 11 types of errors, including word form error(fm), verb phrase phrase error(vb), noun phrase error(np), pronoun(pr), adjective phrase error(aj), adverb phrase error(ad), preposition phrase error(pp), conjunction error(cj), wording error(wd), collocation error(cc), syntax error(sn). The result is achieved through inputting "vb", which stands for the verb phrase phrase error, into the search engine. The total number of verb phrase errors was counted through running "Concordance" of AntConc.

3.2.3. Detagging tool

Because CLEC is a tagged corpus, it is necessary to use a detagging tool to delete all information and tags to analyze students' raw samples.

3.2.4. Judges

The lower grade non-English-major corpus and the higher grade non-English-major corpus in CLEC offers each writing score which was given by a rater (college teacher) on the basis of Principles of Scoring Writing in the College English Test Band Four (CET 4), but two more college English teachers who have been involved in the work of scoring writing for several years and are quite experienced raters were chosen as the judges (raters) to score the writing again in order to ensure the reliability of the evaluation.

Because CLEC is a tagged corpus, the information and annotations in the writing needed to be deleted by means of a detagging tool, the raw compositions were then given to two judges to be scored. The judges were asked to read each composition and rank it on a fifteen-point scale based on the Principles of Scoring Writing in CET4. According to the principles, a holistic method of scoring was employed

which means that the judges' attention was to be focused on the overall effect of the writing rather than on the specific aspects of the writing (such as spelling, choice of words, grammar, etc.). The mean score of each composition, which was received from two teachers' scores and the score offered by corpus, was calculated and recorded as the writing proficiency level of the subjects.

3.2.5. *The statistical softwares*

Excel was employed to calculate the proportion of each type of verb phrase errors and SPSS 21. was used to analyze all verb phrase errors in detail.

Based on the scores that the raters had marked on each composition, the sample compositions were ranked into three levels. Those compositions belonging to the top fifty scores were considered high-level writing. Those compositions belonging to the bottom fifty scores were considered low-level writing, the writing, with the remaining compositions belonging to middle-level writing. Then T-test in SPSS program was run to see which type of verb phrase errors most significantly distinguishes the good writing from the poor writing. The statistical results were explained and analyzed by using relevant acquisition theories. Finally, answers to the research questions were discovered and some conclusions of the present study were drawn.

4. RESULTS AND DISCUSSION

This study adopts quantitative and qualitative methods to analyze the verb phrase errors in college English writing.

4.1. Quantitative Analysis of Verb Phrase Errors

4.1.1. *Verb Phrase Error Distribution*

The descriptive data of the language errors in the table 1 show that there are totally 4586 language errors in the writing sample and every student committed language errors. The table also shows that the number of errors made by each of the sample students is 15.76 on average. The minimum occurrence of the errors is 2, and the maximum occurrence is 44.

In terms of verb phrase errors, there are 554 verb phrase errors which account for 12.08% out of all language errors in the writing sample. Verb phrase errors occur on average 1.91 times in each student' writing sample. For the two extremes of minimum and maximum occurrences, 68 of the students did not make verb phrase errors and one of the students made 11 verb phrase errors.

Table1. *Descriptive Statistics for Verb phrase Errors and Language Errors*

Error type	N	min	max	sum	percentage	mean	std.deviation
verb phrase error	222	0	11	554	12.08%	1.91	1.756
language error	290	2	44	4586	100.00%	15.76	6.973

Note: *N= number of students; sum= total number of errors*

4.1.2. *Differences in Verb Phrase Errors between the High-Score and the Low-Score Writing*

Among the 290 sample writing, the top 50 were labeled high-score writing, the bottom 50 were labeled low-score writing, and the rest were labeled intermediate-score writing. Table 2 shows the difference in verb phrase errors between the high-score and the low-score writing. It can be noticed that there are in total 102 and 66 verb phrase errors in the low-score and high-score writing respectively. The low-score students made an average of 2.04 verb phrase errors, while the high-score students made an average of 1.32. It indicates that low-score students made verb phrase errors twice more than high-score students. Through further ANOVA analysis, it can be seen that there is a strong significant difference in verb phrase errors between the high-score and the low-score writing (F=3.471, p<0.05).

Table2. *Difference between the High-Score and the Low-Score Writing*

Error types	Samples	N	Mean	Std. Deviation	Min	Max	Sum	F(ANOVA)
verb phrase error	LS	50	2.0400	1.77235	.00	7.00	102	3.471*
	IS	190	2.0316	1.81384	.00	11.00	386	
	HS	50	1.3200	1.39152	.00	5.00	66	
language error	Total	290	1.9103	1.75652	.00	11.00	554	

Note: *1. HS=high-score writing; IS=intermediate-score writing; LS=low-score writing*

* means the degree of difference between the groups is significant on the level of 0.05.

4.2. Qualitative Analysis of Verb phrase Errors

Verb phrase errors, including many subtypes, such as errors in agreement, errors in transitivity, errors in voice, errors in mood, etc., occurred 554 times in total. Among them the first three subtypes are most prominent and account for 64% out of verb phrase errors, namely, errors in agreement (214 occurrences), errors in transitivity (104 occurrences), errors in voice (37 occurrences), thus only the three subtypes are explained in detail.

4.2.1. Errors in Transitivity

Verb phrase errors in transitivity are of three types --- taking transitive verb phrases as intransitive ones (47occurences), intransitive verb phrases as transitive ones (34 occurrences) and the wrong use of transitivity of those verb phrases whose transitivity is not word-specific but meaning-specific (23 occurrences). Difference in transitivity of verb phrases between English and Chinese may account for quite a number of verb phrase errors made because of learners' supposition of nonexistent word-for-word equivalents in English. It is known that there are intransitive and transitive verb phrases in both English and Chinese, but it by no means follows an L2-verb phrase-to-L1- counterpart-matching rule. For some verb phrases, they might be intransitive or transitive in both English and Chinese (e.g. *sit* and *zuo*, *learn* and *xuexi*). This seems to require less effort for acquisition. What is worth noticing is that problems arise in the use of English intransitive verb phrases as transitive verb phrases when their Chinese equivalents or near equivalents are transitive. And the reverse is true for English transitive verb phrases. For example in English the verb phrase, depend, is intransitive, while one of its semantic equivalent in Chinese, *yikao*, is transitive. The transitivity of *yikao* here has the potential of resulting in a 'depend + noun phrase' in the Chinese learner's production, for example, never forget to *depend our country* (vi as vt). Such a phenomenon can be explained, it seems, in terms of negative transfer. The following are more examples:

Example 1: They *enjoy* together at conversation, games and amusements. (vt as vi)

Example 2: Children's entertainment is *listening* stories. (vi as vt)

Example 3: Women have full rights to let men *serve for* them.

In example 1, the verb phrase *enjoy* (vt. 0nly) requires the co-occurrence of either a common noun or a reflexive as its object. It is the assumption of the equivalent *xiangle* in Chinese that leads to the learner's use of *enjoy* as an intransitive verb phrase. In example 2, the verb phrase listen is only intransitive in their modern use and yet used transitively in spite of their respective requirement of preposition 'to'. The omission of the preposition is a result of L1 influence because the verb phrase equivalent ting in Chinese is transitive. The verb phrase serve can be either transitive or intransitive. In example 3, the linguistic context requires its use as a transitive verb phrase, but for is inserted between *serve* and them possibly due to his knowledge of the L1 pattern: *wei --- fuwu*.

4.2.2. Errors in Voice

Voice is a universal linguistic phenomenon existing in most languages. While the active voice in English and Chinese is realized both by the agent-predicate-patient order, the passive voice in the languages differs in the restriction of passive marker (*ba* or *bei*) occurrence and verb phrase inflection. Such differences between languages might be the cause of negative transfer found in the Chinese learners' production of passive senesces. For example, *children often criticized* by their parents. It shows an error typical in passive formation, where the passive marker '*ba*' doesn't occur but the preposition '*by*' in the by-phrase possibly serves its function in the learner's mind at the moment of production. Surely the omission of 'ba' in the sentence has much to do with the difference.

4.2.3. Errors in Agreement

Error in agreement refers to error in number agreement with its subject (noun or pronoun). In English, all the finite verb phrases are subject to singular inflection on condition that their grammatical subjects are singular in meaning and take a singular form. In cases of common nouns it follows the general principle of one-position inflection, namely the plural inflection of a countable noun excludes the singular inflection of a finite verb phrase, or to put it another way, the morpheme -s, functioning as an indicator of either a plural noun or a third person singular verb phrase, can, and must, occur only once in a subject–predicate combination, as in: √The students need help;√The student needs help. And a violation of this principle --- with both noun and verb phrase inflection or without inflection at both ---

will certainly result in an error of agreement, for example: *The students needs help*; *The student need help*. In Chinese, however, the syntactic inflection of singularity of verb phrases is totally absent, while that of plurality for nouns exists but is reflected in different ways.

Error in agreement accounts for 38.63% of total verb phrase errors and is most frequently occurring of all subtypes of verb phrase errors. It can be divided into three categories: 1) the verb phrase was unchanged when it should be changed in order to agree with the subject, like examples (4, 5, 6); 2) the verb phrase was changed when it should not be changed in the linguistic contexts, like examples (7, 8); 3) the verb phrase is wrongly changed, like examples (9, 10, 11).

Example 4: It *make/√makes people rich.

Example 5: Human being *have/√has a good living condition.

Example 6: What *make/√makes the great changes in developing country

Example 7: We also *builds /√build hope for the future

Example 8: These changes *does /√do not taken place suddenly.

Example 9: It *changs/√changes very fast.

Example 10: The reasons *is/√are the following.

Example 11: There *were/√is a very low life expectancy.

Through observation of learners' corpus, the first error type is much more common than the second and the third. It was possibly caused by L1 negative transfer because verb phrases in Chinese have no inflection. The second error type is caused by overuse of agreement rule and the third by the learners' incomplete mastery of agreement rule. It is worth noticing that the verb phrase be was often incorrectly changed. The learners got used to *using is* in different linguistic contexts. It indicated that it is difficult for Chinese learners to master different forms of the verb phrase be.

5. CONCLUSION

In sum, the verb phrase errors in the writing sample account for 12.08% of the total language errors. There is a significant difference in verb phrase errors made by college learners of different writing ability. Three main subtypes of verb phrase errors are errors in agreement, errors in transitivity, and errors in voice. And some verb phrase errors are interlingual errors, the others are intralingual.

Errors in students' writing are inevitable, but they are curable or avoidable. It is the teacher's duty to help students trace the root cause of their writing mistakes and then work out remedial measures. On the one hand, errors can provide learners with feedback in the process of concept formation and tell teachers about the effectiveness of their teaching materials and techniques, showing them what parts of the syllabus have been inadequately carried out and need further emphasizing or adjusting. Thereby, teachers should develop a tolerable and positive attitude toward students' errors. On the other hand, teachers should be well aware that an extremely permissive attitude towards errors will lead to fossilization of them, and is thus ultimately unhelpful.

REFERENCES

[1] James, C. *Errors in Language Learning and Use: Exploring Error Analysis*. New York: Addison Wesley Longman Limited, p.1, 1998.

[2] Corder, S.P. The significance of learner's errors. *International Review of Applied Linguistics*, No.4, pp: 161-169, 1967.

[3] Gui, S.C. and Yang, H.Z. *Chinese Learner English Corpus*. Shanghai, Shanghai Foreign Language Education Press, 2003.

[4] Chen W.X. Collocation Errors in the Writings of Chinese Learners of English. *Journal of PLA University of Foreign Languages*, Vol.25, No.1, pp.60-62, 2002.

[5] Li, J.Q. and Cai, J.T. The Misuse of the English Articles in Compositions of Chinese College Students — A corpus-based Study. *Journal of PLA University of Foreign Languages*. Vol. 24(6), pp. 58-62, 2001.

[6] Su H.X. An Analysis of Verb phrase Errors in the Interlanguage of Chinese Learners of English. *Foreign Language Education*. Vol. 23(1), PP. 36-41,2002

[7] He, H.Q. An analysis of lexical errors in non-English-majors' writing: A corpus-based study. *Foreign Language World,* No.3, pp.2–9, 2009.

[8] Xia, L.X. An error analysis of the word class: a case study of Chinese college students. *International Journal of Emerging Technologies in Learning*. Vol.10 (3), pp. 41-45, 2015.

AUTHOR'S BIOGRAPHY

He Huaqing, who was born in Wanzhou, Chongqing city of China in 1970, is a college English teacher in China West Normal University. The main area research is Applied Linguistics.