

A Corpus-based Study of Verb-noun Collocation Errors in Chinese Non-English Majors' Writings

Qiongqiong Zou

College of Foreign Language
Hunan University
Changsha, China

Abstract—Based on the sub-corpora ST3 and ST4 of Chinese Learner English Corpus, this study seeks to investigate the collocation errors made by Chinese non-English majors. The distribution of all six types of collocation errors indicates that Chinese EFL learners have the most difficulty in producing verb-noun collocations. A further qualitative analysis of these verb-noun collocation errors reveals that most verbs and nouns with high rate of collocation errors are those most frequently used ones and that these verb-noun collocation errors can be attributed to various factors, such as first language transfer, misuse of synonyms, overgeneralization and unfamiliarity with semantic prosody. In the light of these findings, some tentative suggestions have been put forward for the teaching and learning of English collocation in China.

Keywords—verb-noun collocation errors; Chinese non-English majors; first language transfer; corpus

I. INTRODUCTION

In recent years, there has been a surge of people around the world becoming learners of English as a foreign language (EFL) in order to meet the needs of intercultural communication. Since vocabulary is the fundamental part of language construction, there is no doubt that the starting point to learn a foreign language is to master a great number of its words. However, despite knowing the lexical meanings of these words, most learners still have great difficulties in using them appropriately in a certain context. A possible explanation may be the fact that words do not exist in isolation but in collocations.

The word “collocation” was first used as a technical term by the British linguist Firth who defined it as “the company that a word keeps” [1], based on which, many other linguists have tried to define collocation from different perspectives since then on. While little consensus on its definition has been reached so far, collocation has attracted increasing attention and has been extensively studied in the field of second language acquisition (SLA) over the past few decades. Moreover, collocation knowledge and skills have already become crucial criteria for measuring the overall language competence of EFL learners nowadays. Although great importance has been attached to English vocabulary teaching in China during the last several decades, it seems that neither teachers nor students have paid much attention to collocations. As a result, Chinese EFL learners tend to show

poor collocation competence and are likely to use more or less inappropriate collocations in various situations ranging from daily oral communication to formal written tasks.

In order to understand the nature of these collocation errors and then to provide some effective suggestions for vocabulary teaching and learning, a great amount of research on collocation has been carried out both abroad and at home. Biskup explored the collocation competence of 34 Polish and 28 German learners of English by asking them to translate some native collocations into English [2]. Bahns and Eldaw conducted an empirical study on 58 German advanced EFL learners by using a translation task and a cloze task in order to examine their competence of English collocations [3]. Gitsaki investigated Greek advanced English learners’ collocation knowledge through three tasks, namely an essay writing test, a blank-filling test and a translation test [4]. Findings revealed that the second language (L2) production was greatly impacted by their first language (L1) and that even advanced EFL learners had great difficulties in producing English verb-noun collocations correctly. With respect to the relevant studies conducted at home, Chen selected more than three hundred compositions written by junior English majors and briefly analyzed their collocation errors. She concluded that there were three major causes for these collocation errors: deficient awareness of collocation, misuse of some synonyms and negative transfer of L1 [5].

The advancement of science and technology has provided various modern approaches for academic studies, one of which is the corpus-based approach. A corpus is a collection of naturally-occurring language data [6]. With the increasing use of corpora in applied linguistics, a new academic field came into being and gradually developed into corpus linguistics, which offers a new approach to SLA research. The Chinese Learner English Corpus (CLEC), the largest language learner corpus in China, has provided Chinese researchers with plenty of authentic data to study the case of Chinese EFL learners. It consists of compositions written by Chinese EFL learners of five levels and thus it is divided into five sub-corpora: senior high school students (ST2), freshmen and sophomores of non-English majors (ST3), juniors and seniors of non-English majors (ST4), freshmen and sophomore of English majors (ST5), juniors and seniors of English majors (ST6) [7]. Based on the sub-corpora ST3 and ST4 of CLEC, the present study attempts to examine the

collocation errors made by Chinese non-English majors and then to explore the possible causes of them, hoping to come up with some tentative suggestions for the teaching and learning of English collocation.

II. RESEARCH QUESTIONS AND METHODS

By adopting a corpus-based approach, the current study aims to address the following three research questions: (1) What are the greatest difficulties that Chinese non-English majors have in producing English collocations? (2) What factors can these collocation errors be attributed to? (3) What pedagogical implications can be drawn for the teaching and learning of English collocation?

Generally speaking, essays in CLEC are manually error-tagged into 61 types of errors, among which the collocation errors (cc) are our research focus. There are six types of collocation errors in CLEC, namely noun-noun collocation errors (cc1), noun-verb collocation errors (cc2), verb-noun collocation errors (cc3), adjective-noun collocation errors (cc4), verb-adverb collocation errors (cc5), and adverb-adjective collocation errors (cc6) [7].

The software tool AntConc 3.4.4 is employed in the current study to collect the data. The main research procedures are as follows: first, extract all six types of collocation errors by using *Concordance* function of AntConc in ST3 and ST4 and then figure out the respective percentage of each error type. Pick out the most typical error type for further analysis; second, observe all the concordance lines containing cc3 errors and select some verbs and nouns which are most likely to be inappropriately used by Chinese EFL learners by using *Collocates* function of AntConc; third, based on these frequently used verb-noun collocation errors, analyze the possible causes of them and then put forward some pedagogical implications.

III. QUANTITATIVE AND QUALITATIVE ANALYSIS OF VERB-NOUN COLLOCATION ERRORS

According to the quantitative analysis, there are altogether 1932 collocation errors in ST3 and ST4. As shown in "Table I", verb-noun collocation error (cc3) makes up the greatest proportion among all six error types, accounting for 57.14% of the total collocation errors, followed by adjective-noun collocation error (14.23%) and noun-verb collocation error (13.82%). It implies that English verb-noun collocation turns out to be the greatest challenge for Chinese non-English majors.

TABLE I. THE DISTRIBUTION OF EACH COLLOCATION ERROR TYPE IN ST3 AND ST4

Error type	Frequency	Percentage
cc1 (n/n collocation)	181	9.37%
cc2 (n/v collocation)	267	13.82%
cc3 (v/n collocation)	1104	57.14%
cc4 (a/n collocation)	275	14.23%
cc5 (v/ad collocation)	83	4.30%
cc6 (ad/a collocation)	22	1.14%
Sum	1932	100%

In order to figure out the features and possible causes of these deviant verb-noun collocations used by Chinese college students, a further qualitative analysis should be conducted. *Collocates* function of AntConc is employed to rank the words which are most likely to be wrongly used. By setting the window span from 5L to 5R, the following verbs and nouns are found to rank top ten in terms of the frequency of wrong use (see "Table II"). Except for the verb "learn", the other four verbs of the top five, i.e. "do", "have", "make", "get", are usually considered as light verbs in linguistics. They have little semantic content of their own when collocating with a noun. "Table II" indicates that Chinese non-English majors encounter great difficulties when it comes to the collocation of these light verbs. "Table III" gives a more detailed description of their collocation errors. There are not such expressions in English as "do practice" "do work" "have health" "have harm" "make success" "get progress". However, we can find their counterparts in Chinese "做练习" "做工作" "拥有健康" "有危害" "获得成功" "取得进步", which suggests that such kind of collocation error, to some extent, results from the literal translation from their mother tongue.

TABLE II. TOP TEN VERBS AND NOUNS IN TERMS OF THE FREQUENCY OF WRONG USE

Verb	Frequency	Noun	Frequency
learn	124	knowledge	117
do	92	society	89
have	76	people	67
make	67	job	51
get	51	commodities	51
work	48	time	48
take	42	water	42
study	29	sex	33
want	26	life	31
know	26	health	29

TABLE III. COLLOCATION ERRORS OF THE FOUR VERBS "DO", "HAVE", "MAKE", AND "GET"

Verb	Collocate
do	practice, problem, work, activity, achievement, dream, point, doctor, scientist...
have	health, harm, illness, medicine, lesson, change, achievement, responsibility...
make	success, communication, life, development, service, practice, trouble, danger...
get	gain, progress, success, influence, achievement, development, harvest, goal...

As for those nouns with high rate of inappropriate use, "knowledge" and "society" are finally picked out for further analysis mainly based on two reasons. For one thing, compared with other nouns in the top-ten list, these two words tend to be more general and less topic-specific since we often speak of them in daily communication. For another, there exist a great variety of words to collocate with "knowledge" and "society". Thus, it is more representative to analyze their high frequency of wrong use. As described in

"Table IV", "knowledge" usually collocates with verbs like "learn" "study" "teach" "widen" by Chinese non-English majors and "society" with "learn" "realize" "touch" "enter" etc. All these wrong verb-noun collocations can also be attributed to the first language because of their Chinese equivalents such as "学习知识" "教知识" "拓宽知识" "接触社会" "步入社会".

TABLE IV. COLLOCATION ERRORS OF THE TWO NOUNS "KNOWLEDGE" AND "SOCIETY"

Noun	Collocate
knowledge	learn, study, teach, know, widen, obtain, increase, benefit, understand, grasp...
society	learn, realize, touch, know (about), enter, step into, go to, live, take part in...

IV. CAUSES OF VERB-NOUN COLLOCATION ERRORS

According to "Table III", "Table IV", and some other verb-noun collocation errors retrieved in ST3 and ST4, we have figured out the following four main causes of verb-noun collocation errors, as summarized in "Table V". It should be pointed out that the possible causes of verb-noun collocation errors made by Chinese non-English majors are quite intricate on account of the fact that most of the errors can be attributed to more than one factor. For example, the inappropriate collocation "get progress" can be possibly due to either the literal translation of its Chinese counterpart "取得进步" or the overgeneralization of the verb "get", which is a light verb often collocating with nouns such as "job", "prize", "gift", etc.

TABLE V. CAUSES OF VERB-NOUN COLLOCATION ERRORS

Cause	Example	Counterpart in Chinese
L1 transfer	learn knowledge touch the society face the life defeat one's weakness	学习知识 接触社会 面对生活 打败某人的弱项
Misuse of the synonyms	grow the medical standard master the knowledge meet difficulty raise the ability	提高医疗标准 掌握知识 遇到困难 提高能力
Overgeneralization	play computer say English get progress take a great improvement	玩电脑 说英语 取得进步 提高
Unfamiliarity with semantic prosody	cause living conditions improved cause the development commit well your business many changes happened	使生活条件改善 使发展 做好你的工作 发生变化

The aforementioned causes of verb-noun collocation errors made by Chinese non-English majors can be generally divided into two categories, namely interlingual factors and intralingual factors, which correspondingly bring about two kinds of collocation errors, i.e. interlingual errors and intralingual errors.

Interlingual errors often result from first language transfer which is defined by Odlin as "the influence resulting from similarities and differences between the target language (TL) and any other language that has been previously acquired" [8]. As we know, language transfer can be essentially classified into positive transfer and negative transfer. The former refers to the transfer which makes the process of L2 learning easier, and always occurs when the L1 and the TL share some similar forms [9]. By contrast, the latter refers to the transfer which makes SLA more difficult and often takes place when there are many differences between the mother tongue and the target language [9]. In the current study, a large number of deviant verb-noun collocations can be ascribed to negative L1 transfer. To be more specific, Chinese EFL learners tend to translate a Chinese collocation into English literally and directly. For instance, they use the collocation "study/learn knowledge" for "学习知识", use "teach knowledge" for "教知识", use "have health" for "拥有健康", use "touch the society" for "接触社会", use "face the life" for "面对生活", etc.

Intralingual errors, on the contrary, usually arise from faulty or partial learning of the target language, instead of from language transfer [10]. Except L1 transfer, the other three factors listed in "Table V" can all be regarded as the causes of intralingual errors. First, misuse of the synonyms: although most Chinese EFL learners presume that there should be a one-to-one mapping relationship between the target language English and their mother tongue Chinese, the actual case is far more complex. One Chinese expression may have a couple of English counterparts, which may share similar meanings but differ in their collocation rules. Therefore, it is possible for Chinese EFL learners to misuse some synonyms. For example, the Chinese word "遇见" has more than one equivalent in English, such as "meet" and "encounter". It is noteworthy that "meet difficulties" is a deviant verb-noun collocation whereas "encounter difficulties" is not. Second, overgeneralization: when the learners have acquired some collocation patterns of certain words in the target language, they are prone to generalizing them and creating some new collocations by themselves, some of which may be incompatible with the collocation rules or the conventional expressions of the target language. Such cases are often considered as overgeneralization, which usually occurs for light verbs, as those in "Table III". Third, unfamiliarity with semantic prosody: Louw defined semantic prosody as "a consistent aura of meaning with which a form is imbued by its collocation" [11]. According to this definition, Stubbs divided semantic prosody into three types: positive prosody, negative prosody and neutral prosody [12]. Positive prosody refers to the favorable semantic aura formed between the node word and its collocation, whereas negative prosody means the unfavorable aura between them. As for neutral prosody, the node word can co-occur with both favorable and unfavorable items, leading to a mixed semantic aura. In this way, if EFL learners are unfamiliar with the semantic prosody of a certain word, they may produce inappropriate verb-noun collocations. This case can be illustrated by the erroneous verb-noun collocations of "cause". It is widely acknowledged that the semantic prosody of "cause" is negative, which indicates that it often

collocates with negative items like “problem”, “damage”, and “death”. Accordingly, collocations like “cause people’s living conditions improved” and “cause the development” are definitely incompatible with the negative semantic prosody of this verb. The reason why Chinese non-English majors produce such wrong expressions is that they may mistakenly consider that the meaning of “cause” is equal to “使” in Chinese. Such cases can also be found in the collocations of “commit”, “happen” and so on.

V. PEDAGOGICAL IMPLICATIONS

Despite the importance of collocation knowledge, the present study has revealed the poor collocation competence of Chinese college students, and more specifically, of Chinese non-English majors. Based on the analysis of verb-noun collocation errors made by Chinese non-English majors and the aforementioned possible causes of these errors, some pedagogical implications can be drawn for English collocation teaching and learning in China.

On the one hand, the fundamental step of teaching verb-noun collocations is to raise the students’ awareness of this language phenomenon. In other words, it is essential for learners to realize the importance of collocations. As displayed in ST3 and ST4, most of the verbs and nouns with high rate of collocation errors are those most frequently used ones, such as “do”, “have”, “make”, “get”, “knowledge”, etc., which indicates that Chinese EFL learners just know the meaning of these words but have no idea of their collocation behaviors. Therefore, during the teaching process, more attention should be paid to collocations, especially to those which are different between Chinese and English. Furthermore, teachers are supposed to compare some English verb-noun collocations with their possible Chinese equivalents in order to make it clear that some equivalents are just grammatically well-formed but pragmatically inappropriate and unacceptable to native English speakers.

On the other hand, vocabulary should be taught in context in the future rather than in isolation. That is to say, mastering an English word means not only knowing its lexical meaning, but also understanding its collocation rules and the specific context where it can be used. It is universally acknowledged that taking English classes is the most typical way to learn English in China, which suggests that Chinese EFL learners lack sufficient and contextualized input of the target language. Therefore, it is understandable that they incline to use some inappropriate and unnatural expressions to convey what they intend to mean. Hence, it is advisable that native corpora should be incorporated into vocabulary teaching so that the authentic context where relevant collocations are properly used will be available to learners, which will help them gain an insight into how to use these words correctly and appropriately in real communication.

VI. CONCLUSION

The current study has explored the verb-noun collocation errors in Chinese non-English majors’ writings by adopting a corpus-based approach. Results demonstrate that among the six types of collocation errors, verb-noun collocation turns

out to be the greatest challenge for Chinese EFL learners. Most verbs and nouns with high rate of collocation errors are those most frequently used ones, which implies that simply knowing the lexical meaning of a certain word does not guarantee that it will be produced correctly when collocating with other words. The underlying causes of these collocation errors are mainly classified into two categories, namely interlingual factors and intralingual factors. The former refers to first language transfer while the latter involves the misuse of synonyms, overgeneralization and unfamiliarity with semantic prosody. Findings of the present study suggest that Chinese EFL learners should pay more attention to collocation knowledge and teachers should integrate native corpora and more authentic context of collocation use into English vocabulary teaching.

REFERENCES

- [1] Firth, J. R. (1957). A synopsis of linguistic theory. In F. R. Palmer (Eds.), *Selected papers for J. R. Firth*. London: Longman.
- [2] Bishup, D. (1992). L1 influence on learners’ renderings of English collocations: A Polish/German empirical study. In P. J. Arnaud & H. B’joint (Eds.), *Vocabulary and Applied Linguistics*. Houndmills: Macmillan.
- [3] Bahns, J., & Eldaw, M. (1993). Should we teach EFL students collocations? *System*, 21(1), 101-114.
- [4] Gitsaki, C. (1999). *Second language and lexical acquisition: A study of the development of collocation knowledge*. Maryland: International Scholars Publications.
- [5] Chen, W. X. (2002). An analysis of collocation errors in Chinese EFL learners’ writings. *Journal of PLA University of Foreign Languages*, 25(1), 60-62.
- [6] Sinclair, J. (1991). *Corpus, concordance, collocation*. Oxford: Oxford University Press.
- [7] Gui, S. C., & Yang, H. Z. (2003). *Chinese learners English corpus*. Shanghai: Shanghai Foreign Language Education Press.
- [8] Odlin, T. (1989). *Language transfer*. Cambridge: Cambridge University Press.
- [9] Ellis, R. (1997). *Second language acquisition*. Oxford: Oxford University Press.
- [10] Richards, J., & Schmidt, R. (2002). *Longman dictionary of language teaching and applied linguistics*. London: Longman.
- [11] Louw, B. (1993). Irony in the text or insincerity in the writer? The diagnostic potential of semantic prosodies. In M. Baker, G. Francis & E. Tognini-Bonelli (Eds.), *Text and technology: In honor of John Sinclair*. Amsterdam: John Benjamins.
- [12] Stubbs, M. (1996). *Text and corpus analysis*. Oxford: Blackwell.