

A Comparative Study on Verb-Noun Collocation Errors of Chinese EFL Learners with Different Language Proficiency

WANG Yi-yin, CHEN Ying-rong

School of Foreign Languages, Zhejiang GongShang University, Hangzhou, China

This study presents a comparative analysis of verb-noun collocation errors made by Chinese EFL learners with different language proficiency. Linguistic data from two sub-corpora of CLEC (Chinese Learner English Corpus, Gui & Yang, 1999) consist of the compositions written freely by Chinese senior high school students (ST2) and the 3rd- and 4th-year English majors of Chinese universities (ST6). The study results show that four types of verb-noun collocation errors are frequently made due to incorrect L1 translation, misuse of relative synonyms, misuse of delexical verbs, and misuse of part of speech; the number of errors decreases with the development of language proficiency, especially errors by incorrect L1 translation and relative synonyms. And for both ST2 and ST6 EFL learners, errors resulted from incorrect L1 translation account for nearly a half with the same high-frequency words. It is also found that ST6 learners' competence in collocation and synonym identification doesn't grow with the increase of their vocabulary, as is shown in the case of overusing delexical verbs in collocation, the same error made by ST2 learners, who have a significantly weaker command of delexical verbs. Implications of the findings are also discussed with the purpose of raising the awareness of collocation pedagogy and enhancing EFL learners' collocation competence.

Keywords: verb-noun collocation errors, CLEC, Chinese EFL learners, a comparative study

Introduction

There is a growing recognition that one of the ways to improve learners' competence in language comprehension is to teach collocations, which can also boost the effectiveness and authenticity of language production (Hashemi et al., 2012; Nuray & Bulent, 2015). Collocation, is defined as "statements of the habitual or customary places of that word" (Firth, 1968, p. 181). Since the 1970s, linguists have ascertained that the linguistic output of native English speakers is not only a single word but also composed of various collocations (Nattinger & Decarrico, 1992). It also appears that collocation is pivotal to L2 acquisition. Unfortunately, it has generally been argued that collocation errors constitute a high percentage of all errors by second language learners, especially verb-noun collocation errors. Researches show that English learners, including high-proficiency learners, are lack of collocational knowledge and underproduce verb-noun collocations (Bahns & Eldaw, 1993; Laufer & Waldman, 2011). Studies done in China also show that Chinese learners of English encounter great

WANG Yi-yin, B.A., School of Foreign Languages, Zhejiang GongShang University.

CHEN Ying-rong (Corresponding author), Associate Professor, School of Foreign Languages, Zhejiang GongShang University.

difficulties in producing correct verb-noun collocations, with improper use of verbs being the main feature of verb-noun collocation errors (Fan, 2004; Zhang & Li, 2004; Zhang & Yang, 2009; Xia et al., 2014; Wang & Li, 2018).

Nevertheless, there remains a paucity of in-depth research that analyzing sub-types of verb-noun collocation errors produced by Chinese EFL learners. Despite certain studies on learners' collocations based on a sizable amount of data from natural production (e.g. Fan, 2004; Zhang & Li, 2004; Zhang & Yang, 2009; Wang & Li, 2018), with findings indicating high percentage of verb errors made by Chinese high-proficiency English learners in their output of verb-noun collocations, the sub-types of verb-noun collocation errors which are problematic for Chinese EFL learners are still unknown. In addition, insufficient attention has been focused on collocation errors made by low-proficiency learners in China, while most previous studies put emphasis on the errors made by high-proficiency learners. What's more, the comparative studies on collocation errors between different levels of language proficiency often rely on frequency comparisons of verb-noun collocation errors, instead of elaborated comparisons of characteristics and categories of verb-noun collocation errors. The existing literature only mentions the subcategories briefly. For instance, Zhang and Yang (2009) merely provided statistical data on verb-noun collocation errors by misuse of delexical verbs, whereas Wang and Li (2018) only presented concrete examples to illustrate Chinese EFL learners' misuse of verbs that are semantically related in verb-noun collocation errors. Likewise, Nesselhauf (2003) indicated that verb errors were the primary source of verb-noun collocations, which frequently happened when a verb was mistaken for another with a similar meaning.

Therefore, our study attempts to further explore the subgroups of verb-noun collocation errors made by low-proficiency and high-proficiency Chinese EFL learners, mainly focusing on the verb errors in the verb-noun collocations. To this end, we have decided to conduct a corpus-based approach, to explore the similarities and differences of verb-noun collocation problems between low-proficiency and high-proficiency learners. Two research questions motivate the study: (a) What are the sub-types of verb errors in verb-noun collocations made by Chinese EFL learners with different language proficiency? (b) What are the similarities and differences in the output of verb-noun collocations by low-proficiency and high-proficiency learners?

Methodology

Research Subjects

The investigation into verb errors in verb-noun collocations of Chinese EFL learners with different language proficiency is based on linguistic data from the sub-corpus ST2 and ST6 of CLEC (Chinese Learner English Corpus), which was edited by Gui Shichun and Yang Huizhong in 1999. The data consist of the compositions written freely by Chinese senior high school students (ST2) and Chinese university students of English majors, mainly in their 3rd and 4th years (ST6). In order to make a comparison of the characteristics of the errors produced by learners with different language proficiency, senior high school students in the ST2 are defined as low-proficiency learners, while senior English major students in the ST6 are regarded as high-proficiency learners. Altogether, 211 verb-noun collocation errors are collected from ST2 with 208,088 tokens in total and 162 errors are extracted from ST6 with the total of 226,106 tokens.

Research Procedures

In the first step, all the verb-noun collocation errors were manually extracted. The error marking system of CLEC was adopted, with which verb-noun collocation errors were commonly identified and tagged by “cc3”, along with a hyphen demonstrating the position of the error and a number determining the context in which the error occurred. Thus, verb-noun collocation errors caused by the wrong choice of verb were sorted out according to the position and context.

The second step was to classify and analyze verb errors. *The Oxford Dictionary* and Britain National Corpus (BNC) were used as references to show the acceptable expressions of the native speakers on one hand, and to help determine the Chinese EFL learners’ errors on the other. Former classifications of verb-noun collocation errors defined by previous researchers (e.g. Nesselhauf, 2003; Zhang & Yang, 2009; Wang & Li, 2018) were taken into account to help this study divide verb errors into the following categories: errors by incorrect L1 translation, misuse of relative synonyms, misuse of delexical verbs and errors in part of speech.

Doing the comparative study between groups was the final step. Considering that the total corpus data of the two sub-corpora of CLEC were not the same, this study converted these data into their standardized frequency counts, which was to suppose that each sub-corpus in CLEC had 20,000 tokens, in order to make the numerical data comparable. For example, the number of errors due to incorrect L1 translation in ST2 was converted into standard data as $58/208088 \times 200000 \approx 55.7$. SPSS Statistics 26 was used to figure out the differences between the frequencies of collocation errors among Chinese EFL learners with different language levels by means of the chi-square test. In addition to the analysis of the data, this study also compared the characteristics of the verb errors of verb-noun collocation produced by the learners, and compared the high-frequency deviant verbs between low-proficiency and high-proficiency learners.

Results and Discussion

Four Types of Errors

As shown in Table 1, four types of verb-noun collocation errors decrease with the improvement of language proficiency. Specifically, the type of verb-noun collocation errors caused by incorrect L1 translation ranks the highest both in ST2 (41.9%) and ST6 (46.1%), accounting for nearly a half in all the verb errors. Meanwhile, both in ST2 and ST6, the frequency of misuse of relative synonyms and delexical verbs accounts for almost a quarter. It can be seen that learners at different levels have fair difficulties with the discrimination of relative synonyms and the use of delexical verbs.

Table 1
Standardized Number and Frequency of Errors in ST2 and ST6

Type of errors	Standardized number in ST2	Frequency	Standardized number in ST6	Frequency
Incorrect L1 translation	55.7	47.9%	31	46.1%
Relative synonyms	31.7	27.3%	16.8	25%
Delxicalized verbs	27.9	24%	19.5	28.9%
Part of speech	1	0.8%	0	0

Incorrect L1 translation

It is stated that when second language learners cannot produce the target language successfully, they tend to apply the rules of L1 to the target language. The consistent rules between the target language and the mother tongue promote second language acquisition by learners, i.e. positive transfer. Conversely, negative transfer leads to difficulties for learners to master authentic expressions of the second language, thus generating linguistic errors (Selinker, 1972). Therefore, Chinese EFL learners are inevitably influenced by their mother tongue in the process of second language acquisition. When they have problems in producing English verb-noun collocations, they often resort to Chinese and directly translate the collocations word by word. However, not all Chinese collocations have translation equivalents in English, so the negative transfer of L1 results in the first type of verb-noun collocation errors. Examples of this type of error are shown in Table 2.

Table 2

Examples of Incorrect L1 Translation

Deviant collocation	Chinese equivalents	Target collocation
eat tea	吃茶	drink tea
learn/study knowledge	学习知识	acquire knowledge
make up one's crime	弥补罪过	atone for crime

As shown in Table 2, these errors overtly violate the grammatical and semantic rules of English. For example, both *learn* and *study* occur frequently in two sub-corpora, collocating with the noun *knowledge*, which is corresponding to *xue xi zhi shi* (学习知识) in Chinese. It is obvious that Chinese EFL learners tend to rely on Chinese explanations and equivalents to choose verbs. This result is also consistent with previous studies (Nesselhauf, 2003; Fan, 2004; Zhang & Yang, 2009). For instance, Fan (2004) investigated the relationship between L1 and verb-noun collocation errors and the result indicated that more than half of the errors were influenced by the participants' mother tongue. Zhang and Yang (2009) also revealed that Chinese English learners with different language proficiency relied considerably on L1. So it is obvious that incorrect literal translation under the influence of L1 is a prominent cause of collocation errors.

The above examples show that when students produce collocations by selecting L1 equivalents, they disregard the restrictions set by the L2 context, leading to inauthentic verb-noun collocation errors. Thus, only when learners acquire vocabulary in English context can they master accurate and reliable lexical meaning and usage. However, the differences between mother tongue and second language have become an obstacle in the process of second language acquisition, and such difficulties exist in all stages of foreign language learning. Nevertheless, chi-square test shows that the difference of error frequency of incorrect L1 translation between ST2 and ST6 has statistical significance ($\chi^2 = 7.772$, $df = 1$, $p = .005$), which indicates that enhancing English proficiency can observably reduce the impact of negative L1 transfer.

Misuse of relative synonyms

Synonyms are a group of words that have the same or nearly the same meanings (Richards et al., 1985, p. 462). Although there are plenty of synonyms in English, expressions identified as synonymous still have significant differences in meaning. Synonyms can also be divided into absolute and relative synonyms. Absolute synonyms are often interchangeable in different contexts, whereas relative synonyms cannot be freely interchanged in terms of collocation and semantics, which account for the majority (Carter, 1998). But in general,

word groups that are synonymous with one another often correspond to only one Chinese equivalent. Therefore, Chinese EFL learners are always confused by relative synonyms and misuse them by substituting the appropriate words in a given context (Nesselhauf, 2003; Wang & Li, 2018). For instance, *wash the rice* is judged unacceptable for the subtle differences between *wash* and *rinse*. Other examples are *kindle the candle* (the correct collocation: *light the candle*), *plant fruits* (the correct collocation: *grow fruits*). Also, learners may ignore the semantic constraints of context. Howarth (1998) pointed out that most of the words are not completely equivalent in semantics and cannot be freely substituted by their synonyms in most contexts. For example, the verb phrase *call on* is more formal and appropriate than the verb *ask* when collocated with a subject *country*. However, Chinese EFL learners are inclined to overlook the complex but subtle differences with regard to contexts, thus confusing the relative synonyms and causing collocation errors. Likewise, the chi-square test also reveals that verb-noun collocation errors which are affected by the confusion of relative synonyms decrease prominently with the improvement of second language proficiency ($\chi^2 = 5.030$, $df = 1$, $p = .025$).

Additionally, ST2 Chinese EFL learners make recurrent errors with basic verbs such as *join* (occurrence frequency: 11) and *go over* (occurrence frequency: 6). Nevertheless, ST6 Chinese EFL learners produce errors related to a variety of relative synonyms and it is more evident that learners have enlarged advanced vocabulary, including *revenge*, *extend*, *heed*, *instruct*, etc. Obviously, high-proficiency learners are rich in their vocabulary to cope with daily communication but their collocation competence doesn't parallel to their vocabulary knowledge in the process of second language acquisition (Bahns, 1993). Therefore, according to the corpus analysis, high-proficiency Chinese EFL learners focus on enlarging their vocabulary without simultaneously acquiring the semantics and collocation knowledge of the words, thus resulting in grammatically and structurally sound but inappropriately collocated errors.

Misuse of delexical verbs

The concept of delexical verbs was first put forward by Sinclair (1990). It refers to high-frequency verbs with little meanings such as *do*, *take*, *make*, *get*, *have*, etc. Delexical verbs do not have specific meanings and their meanings are usually expressed by the words collocated with them (Barfield, 2006). For instance, in the phrase of *take a shower*, it is *shower* that retains its normal meaning while *take* loses its meaning. Verb-noun collocations with delexical verbs are highly arbitrary, and there is no strong semantic motivation, thus, these verbs can be collocated with different verbs to express various meanings, which leads learners to believe that delexical verbs are applicable to a variety of contexts and collocations (Altenberg & Grange, 2001; Zhang & Yang, 2009; Wang & Li, 2018).

Through retrieval and classification analysis, this study finds that learners' misuse of delexical verbs falls into two categories. One is to overuse delexical verbs as substitutes for specific verbs, such as *make one's homeland* (the correct collocation: *develop one's homeland*), *do the problem* (the correct collocation: *solve the problem*), *do crime* (the correct collocation: *commit a crime*), *do punishment* (the correct collocation: *impose punishment*). And the other category is that learners tend to be confused by delexical verbs and interchange them casually, considering that the verbs don't carry significant meanings of their own. Typical examples include *do supper* (the correct collocation: *make supper*), *have sports* (the correct collocation: *do sports*), *take great efforts* (the correct collocation: *make great efforts*), *do contribution* (the correct collocation: *make contribution*).

The frequencies of the two sub-categories are shown in Table 3.

Table 3

Frequencies of Delexical Verb Errors in ST2 and ST6

Type of errors	Standardized number in ST2	Frequency	Standardized number in ST6	Frequency
Overuse delexical verbs	21.1	75.9%	12.4	63.6%
Use delexical verbs interchangeably	6.7	24.1%	7.1	36.4%

As shown in the above table, the errors of overusing delexical verbs decrease slightly with the improvement of language proficiency, whereas the number of errors of using delexical verbs interchangeably doesn't change significantly. As for the frequencies of the two sub-categories, overusing delexical verbs is considerably more than using delexical verbs interchangeably both in ST2 and ST6. It indicates that both low-proficiency and high-proficiency learners have a strong tendency of overusing the verbs when they cannot find appropriate specific verbs. The result is congruent with the finding of Zhang & Yang (2009), which indicated that the overuse of delexical verbs was a major difficulty for Chinese EFL learners, due to learners' misconception that it was safe to use delexical verbs when they were not sure about which specific verb to choose from.

In addition, statistics show that high-proficiency learners still make high frequency errors in the use of basic delexical verbs. Take the simple word "do" as an example, the number of occurrences by low-proficiency Chinese EFL learners is 14, while high-proficiency Chinese EFL learners have lots of errors with *do* (occurrence: 17). Although learners are exposed to some of the delexical verbs and their fixed collocations at the beginning of language learning, they are still unable to use them properly when they reach higher levels. So verb-noun collocations of delexical verbs are difficult for Chinese EFL learners in the process of second language acquisition.

Misuse of part of speech

A unique type of error found in ST2 is the misuse of part of speech. For example:

When it asleep, it likes [cc3, -2] a ball.

According to the context, learners use the preposition *like* as a verb, resulting in an incorrect and confusing collocation. As we all know, part of speech determines the grammatical function of vocabulary to a certain extent, and plays an important role in analyzing and understanding the structure and meaning of sentences. Nevertheless, statistically, this type is extremely rare and only made by low-proficiency learners. Therefore, misuse of part of speech in verb-noun collocation errors is not the majority, but does require some attention in the pursuit of linguistic accuracy.

Implications for Collocation Teaching and Learning

It is widely advocated that raising learners' awareness of collocation learning is deemed to be the foremost work for the teacher (Fan, 2004; Laufer & Waldman, 2011; Wang & Li, 2018). It is obvious that both low-proficiency and high-proficiency Chinese EFL learners have significant problems in using proper verbs in verb-noun collocation, although all types of verb-noun collocation errors mentioned in the paper decrease with the improvement of language proficiency. This indicates that learners overlook the importance of collocations and they are not aware that collocations are arbitrary but, to some extent, have certain restrictions. Therefore, it is important to raise learners' awareness of collocations and to emphasize that collocation plays an important role in the accuracy and authenticity of language output.

It is advised that teachers should help learners identify collocations in authentic materials coupled with a wide range of practice to increase their highly contextualized exposure to and use of collocations. For low-proficiency learners, there is a need for teachers to place more emphasis on collocations and their restricted conditions when teaching simple verbs. Especially for the delexical verbs, the inadequate elicitation and presentation of the teaching materials may leave learners with an incomplete acquisition of vocabulary, resulting in a large number of collocation errors (Xie, 2010). So teaching words together with collocations in text context is of great importance. Meanwhile, high-proficiency learners should also be guided to enlarge their vocabulary, enabling them to not only know verbs' meanings, but also become aware of the appropriate contexts in which the verbs are used and collocated, and identify their relative synonyms as well.

Furthermore, explicit comparative teaching of verb-noun collocations in L1 and L2 is of great significance. This study shows that almost half of the errors are due to interference, or negative transfer from the mother tongue, both for low-proficiency learners (46.1%) and high-proficiency learners (47.9%) of English. As James (2001) pointed out, learners successfully output language by finding equivalents from their mother tongue, and they can benefit from positive transfer from the first language but are also hindered by negative transfer from their mother tongue. Therefore, an explicit comparison of an L2 collocation and its seeming equivalent in L1 can facilitate the clarification of their differences in terms of connotations and usages in contexts, and can be effective to reduce or eliminate collocation errors caused by incorrect translation from L1.

To enhance collocation acquisition, it is also suggested that adequate input from authentic language materials should be provided. Corpus-based learning can be promoted in the classroom to achieve this aim and facilitate L2 collocation acquisition. With the help of corpus, learners are more likely to be exposed to real contexts and conventionalized expressions, and to summarize collocation patterns from the corpus and put them to better use. Some English corpora are widely used today and are beneficial to L2 collocation learning, such as the British National Corpus (BNC) and the Corpus of Contemporary American English (COCA).

Conclusion

Based on the previous studies on verb-noun collocation errors, this study further classifies verb-noun collocation errors made by Chinese EFL learners, focusing exclusively on the verb errors: incorrect L1 translation, misuse of relative synonyms, misuse of delexical verbs, and misuse of part of speech. The errors of delexical verbs can be further categorized as overusing the verbs and using the verbs interchangeably. As the result shows, Chinese English learners' competence in verb-noun collocations has increased with the development of language proficiency. Specifically, the frequencies of errors caused by incorrect L1 translation and confusion of relative synonyms have significantly decreased. Nevertheless, errors caused by incorrect L1 translation remain a prominent problem for both low-proficiency and high-proficiency learners, and also, Chinese EFL learners at different levels have fair difficulties with the discrimination of relative synonyms and the use of delexical verbs.

This study is based on the comparative analysis of existing corpora. In future studies, it may be possible to conduct a longitudinal study and follow a group of L2 learners to see their developmental path of L2 collocation learning. We further recommend delving more carefully into examining the psychological processes that lead to

these errors. So the analysis of the data, combined with interviews and questionnaires, is considered to reveal a more generalized picture of the mental activities of the learners' L2 collocation output.

References

- Altenberg, B., & Granger, S. (2001). The grammatical and lexical patterning of MAKE in native and non-native student writing. *Applied Linguistics*, 22(2), 173-195.
- Bahns, J., & Eldaw, M. (1993). Should we teach EFL students collocations? *System*, 21(1), 101-114.
- Barfield, A. (2006). An exploration of second language collocation knowledge and development (Thesis: University of Swansea).
- Carter, R. (1998). *Vocabulary: Applied linguistic perspectives*. New York: Routledge.
- Deng, Y. C., & Xiao, D. F. (2005). A study of collocations of English delexical verbs by Chinese college English learners. *Foreign Languages and Their Teaching*, (7), 7-10.
- Fan, L. Y. (2004). Error analysis of verb-noun collocation in senior English major students' writing. *Journal of Xi'an International Studies University*, 12(3), 17-20.
- Firth, J. R. (1968). A synopsis of linguistic theory, 1930-1955. In F. R. Palmer (Ed.), *Selected Papers of J. R. Firth (1952-59)* (pp. 168-205). London: Longmans.
- Hashemi, M., et al. (2012). The investigation of collocational errors in university students' writing majoring in English. *Procedia-Social and Behavioral Sciences*, 31(5), 555-558.
- Howarth, P. (1998). Phraseology and second language proficiency. *Applied Linguistics*, 19(1), 24-44.
- James, C. (2001). *Errors in language learning and use: Exploring error analysis*. Beijing: Foreign Language Teaching and Research Press.
- Laufer, B., & Waldman, T. (2011). Verb-noun collocations in second language writing: A corpus analysis of learners' English. *Language Learning*, 61(2), 647-672.
- Nattinger, J. R., & DeCarrico, J. S. (1992). *Lexical phrase and language teaching*. Oxford: Oxford University Press.
- Nesselhauf, N. (2003). The use of collocations by advanced learners of English and some implications for teaching. *Applied Linguistics*, 24(2), 223-242.
- Nuray, G., & Bulent, A. G. (2015). Lexical collocations (verb+noun) across written academic genres in English. *Procedia-Social and Behavioral Science*, 182(5), 433-440.
- Richards, J. C., et al. (1985). *Longman dictionary of applied linguistics*. London: Longman.
- Selinker, L. (1972). Interlanguage. *International Review of Applied Linguistics*, 10(3), 209-231.
- Sinclair, J. (1990). *Collins cobuild English grammar*. London: Collins.
- Wang, W. Y., & Li, X. S. (2018). Investigating the use of verb-noun collocations in L2 writing by advanced learners. *Journal of PLA University of Foreign Languages*, 41(1), 90-98.
- Wolter, B., & Yamashita, J. (2014). Processing collocations in a second language: A case of first language activation? *Applied Psycholinguistics*, 36(5), 1193-1221.
- Xia, L. X., et al. (2014). A corpus-based study on Chinese EFL learners' performance of verb-noun collocations: A case study of "ability". *Foreign Language Education*, 35(1), 68-72.
- Xie, J. C. (2010). A corpus-based comparative study on delexical verbs in EFL coursebooks—A case study of delexical verb "do". *Foreign Language Learning Theory and Practice*, (3), 13-17.
- Zhang, J., & Li, W. Z. (2004). An analysis of errors in the verb-noun collocation pattern in COLEC. *Foreign Language Education*, 25(4), 30-32.
- Zhang, W. Z., & Yang, S. C. (2009). An analysis of V-N collocation errors in CLEC. *Journal of PLA University of Foreign Languages*, 32(2), 39-44.