

A Corpus-Driven Investigation of Chinese English Learners' Performance of Verb-Noun Collocation: A Case Study of *Ability*

Lixin Xia¹

¹Guangdong University of Foreign Studies, China

Correspondence: Lixin Xia, Guangdong University of Foreign Studies, Guangzhou, Guangdong, 510420, China.
E-mail: cdhuiyi@aliyun.com

Received: April 10, 2013 Accepted: May 2, 2013 Online Published: July 4, 2013

doi:10.5539/elt.v6n8p119 URL: <http://dx.doi.org/10.5539/elt.v6n8p119>

Abstract

The paper makes a contrastive study on the performance of verb-noun collocation given by Chinese EFL learners based on the CLEC, ICLE and BNC. First, all the concordance lines with the token "ability" in the CLEC were collected and analyzed. Then, they were tagged manually in order to sort out the sentences in the verb-noun collocation structure. After that, a contrast was made between the data in the three corpora. The study found that there was a distinct gap between the use of verb-noun collocation between English native speakers and Chinese EFL learners, and that Chinese EFL learners also gave less satisfactory performance in verb-noun collocation in terms of correctness and appropriateness.

Keywords: Learner Corpora, English Pedagogical Lexicography, English Language Teaching

1. Introduction

As is widely recognized, the importance of collocation for the learner can hardly be overemphasized, and their performance in using it affects the "naturalness" of their language production directly. Therefore, this study aims to assess Chinese learners' performance of V+N collocation in their English texts based on the Chinese Learner English Corpus (CLEC).

The CLEC is made up of written texts by Chinese learners at different levels. As this study focuses on the collocation errors made by college students, we will select the texts written by college students, which amount to about 800,000 tokens in all.

Collocation errors in the CLEC were tagged as CC1 (N+N), CC2 (N+V), CC3 (V+N), CC4 (Adj+N), CC5 (V+Adv) and CC6 (Adv+Adj). After retrieval, a total of 2,421 collocation errors have been collected, among which the type of CC3 errors reaches 1,352. Next comes CC4 (389), CC2 (300), CC1 (268), CC5 (95) and CC6 (25). The distribution of the six types of collocation errors is shown in Fig. 1.

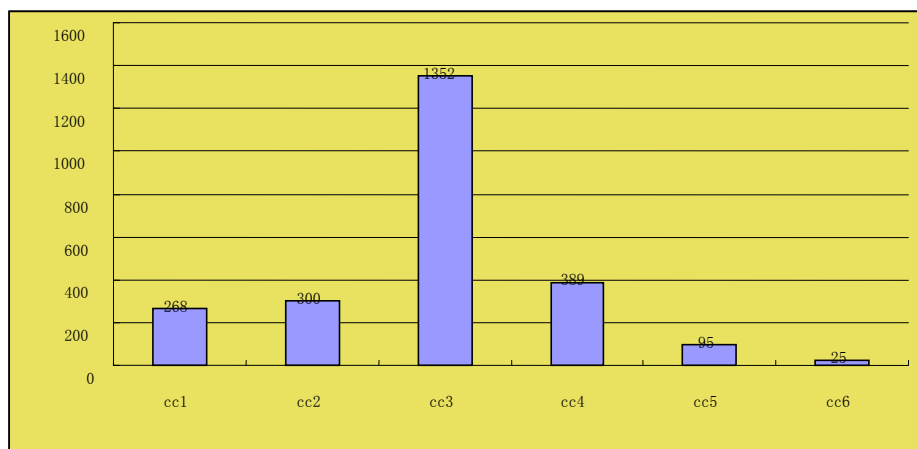


Figure 1. Distribution of the six types of collocation errors

From Fig. 1, we can see that the number of CC3 is even greater than the sum of the other five types of errors. That is to say, Chinese learners are more likely to make collocation errors of V+N. That is the main reason why we focus our study on V+N collocation.

2. V+N Collocation Used by Chinese Learners and English Native Speakers

To have a general picture of Chinese learners' use of the V+N collocation, all the collocation errors tagged as CC3 in the CLEC were retrieved, and the ten lemmas with the greatest frequency were as follows: *learn, knowledge, do, society, have, ability, make, job, get* and *success*.

As *ability* has a high frequency in the CLEC and has not been chosen as a research object in previous studies, we will make a case study of *ability*. All the concordance lines with *ability* in the CLEC and BNC were retrieved, and its frequency in the two corpora was 313 and 10,389 respectively. Since the tokens in the subcorpora of college students in the CLEC and those in the BNC were at a ratio of 800,000 versus 100 million, the rate at which Chinese learners use the word *ability* was four times greater than that at which native English speakers did. So we can reach a conclusion that Chinese learners overused the word *ability* in their written texts.

By using the "Sketch Engine", we got all the V+N collocation of the word *ability* in the BNC, which amounted to 1,795. Table 1 lists the first 51 verbs which collocate with the word *ability* in order of frequency.

Table 1. The collocates of "ability" in the BNC

Collocates	Freq.	Collocates	Freq.	Collocates	Freq.	Collocates	Freq.
show	<u>79</u>	get	<u>28</u>	measure	<u>14</u>	destroy	<u>8</u>
develop	<u>77</u>	lack	<u>27</u>	display	<u>13</u>	learn	<u>8</u>
demonstrate	<u>66</u>	enhance	<u>25</u>	give	<u>13</u>	determine	<u>8</u>
lose	<u>63</u>	reflect	<u>24</u>	involve	<u>12</u>	offer	<u>8</u>
reduce	<u>51</u>	increase	<u>23</u>	read	<u>11</u>	see	<u>8</u>
affect	<u>44</u>	use	<u>22</u>	find	<u>11</u>	inhibit	<u>7</u>
doubt	<u>40</u>	assess	<u>20</u>	underestimate	<u>10</u>	admire	<u>7</u>
require	<u>40</u>	need	<u>19</u>	recognize	<u>10</u>	evolve	<u>7</u>
improve	<u>38</u>	restrict	<u>18</u>	examine	<u>10</u>	undermine	<u>7</u>
limit	<u>35</u>	retain	<u>18</u>	question	<u>9</u>	strengthen	<u>7</u>
include	<u>34</u>	acquire	<u>18</u>	confirm	<u>9</u>	organize	<u>7</u>
test	<u>31</u>	impair	<u>17</u>	reveal	<u>9</u>	match	<u>7</u>
possess	<u>29</u>	prove	<u>17</u>	describe	<u>9</u>	exercise	<u>7</u>

From Table 1, we can see that *ability* is usually used together with the verb *show, develop, demonstrate, lose, reduce*, etc. They can be classified into the following groups according to their literal meanings: *develop/evolve ability, possess/retain ability, show/ demonstrate/ display/ reveal/reflect ability, improve/ enhance/ strengthen/ increase ability, learn/acquire/get/find/see ability, admire/recognize ability, organize ability, use/exercise ability, prove/confirm ability, include/involve ability, read/describe/match ability, require/need ability, give/offer ability, test/ assess/ measure/examine ability, affect ability, doubt/question ability, underestimate/ determine ability, lose/lack/reduce/undermine ability, limit/restrict/inhibit ability, impair/destroy ability*.

Then, all the concordance lines with the word *ability* were retrieved from the CLEC. After tagged manually, we got 212 instances of the V+N collocation, among which 48 verbs were used together with the word *ability* as shown in Table 2.

Table 2. The collocates of “ability” in the CLEC

Collocates	Freq.	Collocates	Freq.	Collocates	Freq.	Collocates	Freq.
have	61	use	3	connect	1	justify	1
improve	36	acquire	2	demonstrate	1	learn	1
develop	14	forge	2	deny	1	maintain	1
show	14	stand for	2	doubt	1	neglect	1
find	8	strengthen	2	enforce	1	perform	1
train	8	test	2	enhance	1	promote	1
need	6	accumulate	1	enrich	1	raise	1
prove	6	adapt	1	examine	1	resort to	1
lose	5	appreciate	1	exert	1	reveal	1
cultivate	4	block	1	form	1	stimulate	1
display	4	bring about	1	gain	1	surpass	1
practice	3	build	1	indicate	1	value	1

We grouped them in the same way and got the following groups: *have/maintain* ability, *develop/cultivate/train* ability, *show/demonstrate/display/reveal/indicate/stand for* ability, *improve/enhance/strengthen/raise/promote* ability, *appreciate/value* ability, *learn/acquire/gain/find/accumulate/enrich/stimulate* ability, *use/exert/ perform/ practice/ enforce/resort to* ability, *forge/build/form/bring about* ability, *adapt* ability, *prove/justify* ability, *need* ability, *test/examine* ability, *doubt* ability, *lose* ability, *block* ability, *deny/neglect* ability, *surpass* ability, *connect* ability.

Based on the above data, we might be able to draw the following conclusions:

First of all, Chinese learners tend to overuse some collocation structure. For example, while the frequency of the first ten collocations of *ability* in the BNC is only from 0.35 to 0.79 in one million, it is up to 5 to 18 in one million in Chinese learners' written texts. That is to say, it is from ten to a hundred times more chances for Chinese learners to use these ten collocations in their writings than for English native speakers do. Besides, the collocation “possess ... ability” occurs 29 times in the BNC, but none in the CLEC. The reason for this might be that Chinese learners have a limited size of vocabulary. Therefore, they have a tendency to choose the words and expressions that they feel safe to use. The collocation “have... ability” is an example of this, which was overused by Chinese learners in the CLEC. Similarly, the collocation “cultivate ... ability” appears four times in the texts in the CLEC, but only once in the BNC.

Next, compared with the phenomenon of overusing some collocations, the cases of underusing some collocations can be frequently found in Chinese learners' written texts. Sometimes the state of underuse is even more remarkable than that of overuse. The case of “possess ... ability” discussed above is an example of this. Moreover, Chinese learners are found to be particularly prone to underuse those collocation structures with negative semantic prosody. For example, English native speakers use such collocations as “*lose/lack/reduce/undermine* ability”, “*limit/ restrict/ inhibit* ability” and “*impair/destroy* ability”, etc. with a frequency of 63, 27, 51, 7, 35, 18, 7, 17 and 8 respectively for the collocates of *lose, lack, reduce, undermine, limit, restrict, inhibit, impair* and *destroy* in the BNC. But in the CLEC, only one collation with negative meaning was found, i.e. “lose ability”. This shows that Chinese learners favor the use of collocations with positive semantic prosody, which might be caused by their negligence in acquiring collocations with negative semantic prosody.

Thirdly, the main causes for collocation errors are mother tongue transfer and intralingual transfer. Among the 48 V+N collocations of *ability* in the CLEC, 16 of them didn't appear in the BNC. They were “*train/ forge/ build/ form/ accumulate/ enrich/ raise/ enforce/ perform/deny/neglect/ justify/ practice/ adapt/ surpass/ connect* ability”. Below are the instances of the 16 collocations excerpted from the CLEC:

- 1) m1,-] for me. I think it not only train my **ability** [cc3,-] to adapt [vp1,-s] to
- 2) rk. [sn2,s] And so, they can forge their **ability** constantly, to adopt [wd3,-3]

- 3) ur problems into action, and build the **ability** to cope with many practice pro
- 4) pl,-s] that a lot of practice can form a **ability** of doing something. For exampl
- 5) e should accumulate knowledge and **ability** [cc3,4-] in usual time [np6
- 6) a good job they have to enrich their **ability** and have some special skill. I can
- 7) tant, as to [pp1,1-4] our people raise **ability** of identify.[sn2,s] So [wd5,-] the
- 8) should be preparing to enforce their **ability** in order to find a good job in the
- 9) le and how to perform [cc3,-2] your **ability** well may be hard to realize
- 10) More [wd3,s,-] I do not deny their **ability** to become rich. But I would not
- 11) for high scores and neglect other **ability** [np6,-] and actions [wd7,-] . In our
- 12) the teachers to justify [cc3,-4] our **ability** and personality. Then we cou
- 13) I think I should practice all kinds of **ability** [np3,-] . Then when I enter
- 14) e same time, adapt to the progress of **ability**. [sn8,s] They can get s
- 15) ce pollutants discharged surpass the **ability** of natural purification, these places
- 16) t seems that we only connect our **ability** or cleverness with our theory examina

In the above examples, Ex. 1, Ex. 2 and Ex. 7-9 were the direct result of using a Chinese collocation. Thus, we had the collocations of “train/ forge/ build/form/ raise/ enforce/ perform...ability”. Ex. 5, Ex. 6 and Ex. 12 were the result of intralingual transfer. In Ex. 5, the predicate verb *accumulate* was followed by two objects, the first of which was correct, but the last of which is not natural. In Ex. 6, *imagination* and *experience* were possible collocates of the verb *enrich*, but *ability* was not. In Ex. 12, the learner might intend to express the meaning of “prove our ability”, but he used *justify* instead. For the rest of the examples, they should not be regarded as wrong collocations as they didn’t violate the collocation rule.

Last but not the least, some forms of the V+N collocation listed above were not tagged as a collocation error of CC3 in the CLEC. Only Ex. 1, Ex. 5, Ex. 9 and Ex. 12 were tagged as CC3. However, some collocations used by English native speakers were listed as collocation errors in the CLEC as follows:

- 17) ause they have many chances to **show** their ability [cc3,2-]. They also think ou
- 18) iving better, the person can **display** his ability, [cc3,2-] wise [wd2,s-] and suc
- 19) the one you like best and can **value** your ability. [cc3, 2-] Everything is decid
- 20) d he been so lonely. He was fired and **took** [cc3,-3] great effort to find a new j
- 21) s to crime easily. Usually, we **give** [cc3,-2] great sympathy to those criminals

This might be caused by the negligence of the corpus builders. Although this may decrease the accuracy of the corpus data, it won’t have much influence on the analysis of the characteristics of collocation errors made by Chinese learners.

3. A Contrastive Study on the Collocation Used by Chinese Learners and International Learners

As the ICLE is an untagged corpus, we selected the written texts by English learners in Spain, France, Netherlands, Finland and Czechoslovakia randomly, which added up to about 800,000 tokens. Therefore, the amount of data chosen from the two learner corpora and the register of the written texts were comparable, which guarantees the validity of the comparison.

Table 3. The collocates of “ability” in the ICLE

Collocates	Freq.	Collocates	Freq.	Collocates	Freq.	Collocates	Freq.
have	14	demonstrate	1	demand	1	omit	1
Lose	10	impair	1	descapacite	1	show	1
Use	3	reduce	1	possess	1	support	1
Prove	2	create	1	retain	1	develop	1
need	2	decrease	1	undermine	1		

After retrieval, we got 120 concordance lines with *ability* in the ICLE, out of which we sorted out 19 V+N collocation by tagging manually. The most commonly used verbs were *have*, *lose*, *use*, *prove*, and *need* as shown in Table 3. Among the 19 collocations, only two of them were not correct collocations. They were “descapacite ... ability” and “omit ... ability” as listed below:

22) As a result, there are many consequences that television has occasioned like: First, television has destroyed the Communication that there was in homes and with friends. Second, has descapacited the ability of imagination because there is not reading

23) It should focus on the most important and fundamental soldier 's skills as handling with a gun, being able to use gas mask , etc. We cannot omit also their ability to work in groups, comprehend the military rules

In Ex. 22, the learner intended to express the meaning of “destroy the ability”, but he used a word that didn’t exist in English. In Ex. 23, the correct form should be “neglect to develop their ability”.

Compared with Chinese learners, the rate at which international learners used the V+N collocation of *ability* was almost the same as Chinese learners. However, their production of the collocation was more natural and correct. In the ICLE, only two errors were found. However, in the CLEC more errors were detected and some of them appeared repeatedly. For example, there were 10 instances of the collocation “train ability” in the CLEC. Therefore, Chinese learners have greater chances of making collocation errors than international learners.

Moreover, international learners behaved in a more native-like way in terms of collocation structure. They frequently used such collocations as “have/ lose/ use/ prove/ possess/ need/ reduce/ demonstrate...ability”, which also had a high frequency in the BNC. In addition, international learners used more collocations with a negative semantic prosody than Chinese learners did. For example, 4 out of the 19 collocations used by international learners had a negative semantic meaning; and they are “lose/reduce/impair/ undermine/decrease...ability”. However, Chinese learners only used one such collocation: lose... ability. The difference reflects the gap between Chinese learners and international learners in using collocations.

4. Conclusions

The V+N collocation of *ability* from the BNC, CLEC, and ICLE was collected and analyzed. The major findings are summarized as below:

First, collocation errors are common in Chinese learners’ writings, especially the V+N collocation. Among the six types of collocation errors, the number of CC3, i.e. V+N collocation errors, is the greatest. This may suggest that Chinese learners have difficulty in acquiring and using the V+N collocation.

Secondly, a distinct difference can be spotted between Chinese learners and English native speakers in using the collocation. Due to the limit of their vocabulary, Chinese learners tend to overuse the collocations that have an equivalent in Chinese, and underuse the collocations with a negative semantic prosody.

Finally, a substantial gap can be found between Chinese learners and international learners with regard to correctness and naturalness of their collocation. In other words, international learners produce more natural collocation than Chinese learners.

References

- Corder, S. P. (1981). *Error analysis and interlanguage*. Oxford: Oxford University Press.
- Ellis, R. (2008). *The Study of Second Language Acquisition* (2nd ed.). Oxford University Press.
- Gui, Shichun, & Yang, Huizhong. (2002). *Chinese Learner English Corpus*. Shanghai: Shanghai Foreign Language Education Press.
- Gui, Shichun. (2005). An analysis of speech errors of Chinese learner. In Yang Huizhong, Gui Shichun, & Yang Dafu (Eds.), *An Analysis of Chinese Learner's Interlanguage based on CLEC* (pp. 1-109). Shanghai: Shanghai Foreign Language Education Press.
- Nemser, W. (1971). Approximative systems of foreign language learners. *International Review of Applied Linguistics*, 1971(9), 115-123.
- Nesselhauf, N. (2003). The Use of Collocations by Advanced Learners of English and Some Implications for Teaching. *Applied Linguistics*, 2003(2), 223-242.
- Selinker, L. (1992). *Rediscovering Interlanguage*. London & New York: Longman.
- Selinker, L., & Lakshmanan, U. (1992). Language transfer and fossilization: The multiple effects principle. In Gass, S. M., & Selinker, L. (Eds.), *Language Transfer in Language Learning* (pp. 197-216).

Amsterdam/Philadelphia: John Benjamins.

Sinclair, J. (1991). *Corpus, concordance, and collocation*. Oxford: Oxford University Press.

Wang, Haihua, & Zhou, Xiujuan. (2009). A study on Chinese learners' development of verb + noun collocation behavior: a corpus-driven methodology. *Waiyu Xuekan*, 2009(6), 59-62.

Copyrights

Copyright for this article is retained by the author(s), with first publication rights granted to the journal.

This is an open-access article distributed under the terms and conditions of the Creative Commons Attribution license (<http://creativecommons.org/licenses/by/3.0/>).