

Parallel Corpora and Translation Teaching

Jingang Bai

School of Foreign Languages, Chifeng University, Chifeng, China

1461942076@qq.com

Keywords: Translation Teaching; Parallel Corpora; Collocation; Translation learning

Abstract. In recent years, the development of corpora-based study provides language learners and teachers with great opportunities to have access to parallel corpora in their translation teaching and learning. In this paper, the author tries to explore whether subjects can solve the practical problem of finding proper collocation with the help of parallel corpora in the process of translation teaching.

Introduction

A corpus can be defined in terms of both its size and its content. A corpus is a “collection of pieces of language that are selected and ordered according to the explicit linguistic criteria in order to be used as the sample of the language” A corpus is “a computerized collection of authentic texts amenable to automatic or semiautomatic processing or analysis.” and “the text are selected according to explicit criteria in order to capture the regularities of a language, a language variety or a sub-language” So there is no doubt that working with corpora can revolutionize the way we teach and learn. Parallel corpora have been widely used in language teaching and learning, especially in writing and speaking, but seldom applied into translation teaching and learning. Parallel corpora have much potential in translation practice. Parallel corpora can provide translators with the insights into strategies employed by past translators in dealing with translation problems. Therefore the author conducted an experiment to investigate that parallel corpora can be useful in the process of translation to help subjects to find proper collocations to improve the quality of their translation versions.

Objectives and Subjects

In translation practice, making natural collocates is one of the difficult problems for subjects in the process of translation. In the experiment, the author mainly focuses on the translation of typical collocations that the students translate with difficulty since parallel corpora can provide collocation information for students to learn especially about adjectives that collocate with nouns and to see which words collocate with the verbs.

The general objective of the paper is to testify if subjects can perform better with the help of parallel corpora than those who translate with conventional dictionaries when two groups of subjects working with the same translation material. Specifically, the research is to testify if parallel corpora can help subjects find more proper collocations for translation candidates than conventional dictionaries. Subjects are 80 sophomores from the computer department of Chifeng University, who are divided into two groups in the experiment. One group of subjects is working with parallel corpora; the other is working with conventional dictionaries. The reasons for choosing those students are as following: the subjects are from computer department, so they have the basic computer skills which are necessary to operate parallel corpora; they are going to attend the examination of CET4. C-E translation is one part of it; as sophomore, they have got the basic translation skills to work with identical translation materials with the help of translation tools.

Test Material and Tools

Test materials are sentences translations and paragraphs translations selected from CET 4 Model test. PACCEL (Parallel Corpus of Chinese EFL learners) will be used in the experiment, which is

the first large scale Chinese-English parallel corpora of Chinese English foreign language learners in China. The parallel corpus has two sub-storehouses: (PACCEL-S) is for spoken learning; (PACCEL—W) is for written learning. The contents of the corpus are juniors' or seniors' translation tests and exercises of English majors from eighteen colleges and universities. In the experiment,(PACCEL—W) contains about sixteen million words with texts in Chinese and their translations in English, and the degree and the scale well meet the need of solving the translation problems in College English Test 4.

In addition, several online parallel corpora are also used in the experiment. Parallel corpora in CQPweb: *The Babel English-Chinese Parallel Corpus* with 327 English articles and their translations in Mandarin Chinese from *World of English* between October 2000 and February 2001 and *Time* from September 2000 to January 2001. The corpus is available on <http://124.193.83.252/cqp/>. Another English-Chinese online parallel corpus is available at <http://www.luweixmu.com/>, which contains 215,713 sentences including 3,290,670 English words and 5,370,429 Chinese characters, is powered by professor Lu Wei from Overseas Education College of Xiamen University.

In the experiment, concordance Paraconc works as the corpus analyzing tool. It is a bilingual or multilingual concordance designed and produced by Professor Michael Barlow in University of Auckland, which can be used in contrastive analyses, language learning and teaching, translation studies and training and so on.

Procedure of Experiment

The experiment lasts about 8 weeks from April 2015 to June 2015, which mainly included two phases: preparation and implementation phase.

Preparation

Since the experiment is conducted to compare and contrast two groups of subjects' performance with different translation tools to translate the same material, the author prepare two groups of subjects, conventional C-E dictionaries and several parallel corpora as well as translation materials.

Subjects in the control group with conventional dictionaries will work in a traditional way that after subjects finish the translation tasks in class or after class, they are required to bring them back to discuss in class. Finally, the author presents the best translation to the class.

Parallel corpus is a kind of computer software, firstly, the author would teach subjects to operate corpora and learn about the corpora analysis tools and how to extract transnationally-relevant information from corpora. On this basis, subjects in experimental group are required to translate the same materials with parallel corpora.

Implementation

Both subjects using conventional dictionaries and parallel corpora are required to translate eight pieces of C-E translation materials, once a week as assignment. Then the author compares the translation versions from two groups and make conclusion. In the experiment, the author mainly compares the words translation and phrases translation to testify the function of parallel corpora in translation practice. That is parallel corpora can help subjects solve the problem of collocation in the process of translation so that subjects can avoid making mistakes and get translation versions with higher- quality.

Results and Analysis

Collocations are defined as “words that appear together with greater than random probability”. That is to say, if the words typically co-occur, they are considered to be collocates. In translation practice, making natural collocates is one of the difficult problems for subjects in the process of translation.

Bilingual dictionaries can only provide the individual equivalent. It is difficult for subjects to decide whether translation candidates collocate with each other. In this aspect, parallel corpora can also be useful translation aids for subjects to solve the problem. It can provide collocation information for subjects to learn especially about adjectives that collocate with nouns and to see which words collocate with the verbs and so on. For example, for the translation of the phrase “anquan shiwu” in the sentence , it is easy to translate “anquan” as “safe, safety, secure, security” and translate “shiwu” as “matters, events, issues, affairs, incident, work”. According to dictionaries, subjects can also find some equivalents such as “safety problems, security matters, security issues, security events and so on”, but they need to assess which collocation sounds more natural in target language, which is the difficult problem that subjects can’t solve only with the help of dictionaries. Parallel corpora play a considerable role in providing students with collocation information. Subjects who translate with corpora are more likely to find proper collocation “security affairs” by tagging “anquan; shiwu” in Keywords List, because in parallel corpora, there are parallel samples containing the translation of “anquan shiwu” used by past translators.

Taking another phrase “zhiding guizhang tiaoli” as an example, for subjects, it is easier to translate “guizhang” and “tiaoli” as “rules” and “regulations”. The problem is which translation of “zhiding” collocates with “rules” and “regulations”. For the translation of “zhiding”, subjects may think of “make”, “draw up”, “Write”, “lie down”, “establish”, “formulate” and so on, according to dictionaries, but dictionaries can’t provide information on appropriateness of translation, and dictionaries can hardly provide the same context containing the phrase “zhiding guizhang tiaoli”. Through parallel corpora, subjects can find accurate collocation by tagging “zhiding; guizhang” and “zhiding; tiaoli” into Keywords List, subjects can find proper translation “formulate rules and regulations” through observing the parallel sample sentences containing “zhiding; guizhang” and “zhiding; tiaoli”. In addition, parallel corpora can also be used to deal with the words which have similarities but have many differences in actual use in the process of translation. Since different language has different collocation patterning, parallel corpora prove to be valuable resources to offer a multitude of collocation varieties and give information on selecting the optional collocation. Next two tables are about the translations of “anquan shiwu” and “zhiding guizhang tiaoli”:

Table 1 Translation of “anquan shiwu”

anquan shiwu	The experimental group (With corpora)		The control group (Without corpora)	
	figure	percentage	figure	percentage
Acceptable expressions				
security affairs	20	50	12	30
security matters	10	25	2	5
security issues	6	15	4	10
security	4	1	0	0
Total	40	100	18	45
Unacceptable expressions				
safety affairs	0		7	17.5
safety matters	0		5	12.5
safety issues	0		4	10
security work	0		1	2.5
safety work	0		3	7.5
safe event	0		2	5
Total			22	55

In this case, all the subjects in the experimental group get the acceptable translations of “anquan shiwu”, such as “security affairs”, “security matters”, “security issues” and “security”, and among which 50% of subjects find the native translation “security affairs”, 25% “security matters”, 15% “security issues” and 1% “security work”.

“security” with the help of parallel corpora, however, in the control group, without corpora, only 45% of subjects’ translations are acceptable, among which 30% of subjects finds the native translation of “security affairs”, but there are still 55% of subjects producing wrong collocations, such as, “safety affairs”, “safety matters”, “safety issues”, “safety work”, “security work” and “safe event”, all of which seem right, but in fact, they don’t collocate with each other.

Table 2 Translation of “zhiding guizhang tiaoli”

Zhiding guizhang tiaoli	The experimental group (With corpora)		The control group (Without corpora)	
	figure	percentage	figure	percentage
Acceptable expressions				
formulate rules and regulations	30	75	6	15
Unacceptable expressions				
make rules and regulations	3	7.5	14	35
establish rules and regulations	7	17.5	10	25
set up rules and regulations	0	0	3	7.5
draw up rules and regulations	0	0	4	1
lay down rules and regulations	0	0	2	5
draft rules and regulations	0	0	1	2.5
Total	10	25	34	85

This is the typical phrase which indicates the problem the subjects usually encounter in choosing the appropriate verbs to collocate with nouns in the process of translation. Students are easy to think of “make”, “formulate”, “establish”, “set up”, “draw up”, “lay down” and “draft”, while translating the phrase “zhiding”, because in their understanding, all of them have the meaning of “zhiding”. But it is hard for subjects to assess the translation candidates and they don’t know which one is the best to collocate with “rules and regulations”. So in the control group, there appear various translations. In the experimental group, 75% percent of subjects get the correct translation “formulate the rule and regulations” with the help of parallel corpora, while in the control group, only 15% of subjects get the correct collocation, which may imply that parallel corpora not only provide students with collocation information that subjects can see how the Chinese lexical items transferred into English and at the same time with the parallel samples subjects can assess if their choices are acceptable or not.

Findings and Conclusion

Through the application of parallel corpora as translation tools to subjects, the author finds that parallel corpora can provide subjects with collocation information which can help them produce more natural target texts out of their mother tongue, which can be proved by the translation of the two phrases “anquan shiwu” and “zhiding guizhang tiaoli”. The translation of “anquan shiwu” proves that parallel corpora can help subjects identify the different use of nouns in typical expressions so that they can find the proper nouns which collocate with each other in translation; the translation of “zhiding guizhang tiaoli” shows another significant effect of parallel corpora that they can provide perfect translations for expressions in which verbs collocate with nouns, which is

the typical problem that subjects always meet in the translation.

So it can be concluded that Parallel corpora provide an effective approach for subjects to solve the problem of collocation. The author asks subjects to do many related translations and chooses to compare the translation of “anquan shiwu” and “zhiding guizhang zhidu” to show the resourcefulness of parallel corpora in helping subjects to assess and evaluate their choices in finding proper collocations.

References

- [1] Pearson, Jennifer. *Terms in Context* [M]. Amsterdam/Philadelphia: John Benjamin Publishing Company. 1998.
- [2] Tognini-Bonelli, Elena. *Corpora Linguistics at work* [M]. Amsterdam/Philadelphia: John Benjamin Publishing Company. 2001.
- [3] Teubert, Wolfgang / Cermakova, Anna. Directions in Corpus Linguistics. In Halliday, Michael A. K. et al. (eds) *Lexicology and Corpus Linguistics. An introduction* [C]. London: MPG Book Ltd, 1992: 113-167.
- [4] Bowker, Lynne / Pearson, Jennifer. *Working with Specialized Language: A Parallel guide to using corpora* [M]. London/New York: Routledge. 2002.
- [5] McEnery, Tony / Wilson. *Corpora Linguistics: An introduction* [M]. Edingurbergh: Edingurbergh University Press. 2001.
- [6] Teubert, Wolfgang / Cermakova, Anna. Directions in Corpus Linguistics. In Halliday, Michael A. K. et al. (eds) *Lexicology and Corpus Linguistics. An introduction* [C]. London: MPG Book Ltd, 1992: 113-167.
- [7] Zanettin, Federico. Bilingual comparable corpora and the training of translators [J]. *Meta* 43/4, 1998: 616-630.
- [8] Leech, G. Introducing Corpus Annotation [J]. In Garside, R., Leech, G., and Tony McEnery (Eds.) *Corpus Annotation*. London: Longman, 1992: 106.
- [9] Machniewski, Maciej. Analyzing and teaching translation through corpora: *Lexical convention and lexical use*. *Poznan Studies in Contemporary Linguistics* [C]. 2006: 237-255.
- [10] Monzo, Esther. Corpus-based Teaching: The Use of Original and Translated Texts in the training of legal translators [J]. *Translation Journal*. 7/4. Retrieved May 2, 2008, from: <<http://accurapid.com/journal/26edu.htm>>.