

A Corpus-Based Study on English Profiles of Chinese and US Companies in Electronic and Information Industry

Lina Mo

Nanfeng College of Sun Yat-sen University, Guangzhou, 510970, China

penseatoi@163.com

Keywords: Corpus, English profiles, Companies, Translation.

Abstract. With the development of computer science, corpus has served as a new approach for translation studies and obtained a growing interest from researchers and scholars in recent decades. With the aid of AntConc, Claws, Wordsmith and other corpus tools, this study attempts to compare the English profiles of 30 top Chinese companies and 30 U.S. companies in electronic and information industry and explore their language features in terms of word frequency, key-words, lexical diversity and density, standardized TTR, mean sentence length and passage length. Through the corpus-based analysis, it is found that compared with the English profiles of Chinese companies, those of the U.S. companies embody stronger customer orientation with more concise and readable texts in friendlier language style. These findings have significant implications for the translation of profiles of Chinese companies.

Introduction

In recent decades, an increasing number of Chinese companies have ventured into the global market. In order to explore business opportunities, many Chinese companies have set up an English or multi-language website, the contents of which mainly include a company profile, products and services, contact information, etc. For Chinese companies, an English company profile is the company's resume in the international market. However, a close look at the Chinese and English websites of a number of Chinese companies reveals that most of the English company profiles are translated literally from the Chinese versions without considering the cultural differences and the expectations of the target readers. The study therefore attempts to compare the English profiles of 30 top Chinese companies and 30 U.S. companies in the electronic and information industry and explore their similarities and differences in language features in the hope that the findings of the comparative study can have certain implications for the translation of Chinese company profiles.

Company Profiles and Related Translation Studies

Ann Hackett in "How to Develop an Effective Company Profile and Why" defines company profiles as a 'resume' with in-depth information about an organization, targeting to establish the credibility with the market the company serves. The company profile is important because it helps the potential customers to understand the company's business as well as its approach, unique strengths, and relevant experience (quoted in Li Quandong, 2013:3). A company profile usually includes an overview of the history, current status of business and future goals of a company. It is often the first place that the potential clients and business partners of the company will navigate to on its website. For Chinese companies, an English company profile is an essential part of its publicity materials. A well-translated company profile can reinforce the corporate image and bring business opportunities while a poor one may lead to misunderstandings and confusions. So far researches on translation of Chinese company profiles have been mostly put in the broader field of translation of publicity materials and carried out from the perspectives of text typology theory, the skopos theory and reception theory. In China, researches on translation of publicity materials started about a decade ago. In the article "Adhering to the Principle of Three Closeness and Handling Difficulties in Translating Overseas-Targeted Publicity Materials" (Huang Youyi, 2004), Chinese Scholar Huang Youyi put forward the principle of three closeness, i.e., close to the reality of China's

development, close to foreign readers’ demand for information about China, and close to the foreign readers’ way of thinking, which provided guidelines for the practice of publicity translation.

In recent decades, the combination of computer and translation studies has given rise to the corpus-based translation studies and provided a new approach for translation studies. Corpus-based translation studies originated from the traditional thought on linguistics represented by R. Firth, M.A.K. Halliday and John Sinclair (Stubbs, 1993:8-13; Laviosa, 2002:5). In 1993, the British scholar Mona Baker, in her article “Corpus Linguistics and Translation Studies: Implications and Applications”, argues that “translation studies has reached a stage in its development as a discipline when it is both ready for and needs the techniques and methodology of corpus linguistics in order to make a major leap from prescriptive to descriptive statements, from methodologizing to proper theorizing, and from individual and fragmented pieces of research to powerful generalizations”. The present study on translation of company profiles is a corpus-based study, which involves creating two corpora, analyzing the similarities and differences between them with the aid of various corpus tools, and finally summarizing the implications for the translation of Chinese company profiles.

Research Methodology

Corpus Collection

The English profiles of the top 30 companies (see Table 1) from “China’s Top 100 Enterprises in Electronic and Information Industry” are collected to build the English profile corpus of Chinese companies (hereinafter referred to as the EPCCCs). In the meanwhile, the English profiles of 30 well-known American companies (see Table 2) in the same industry mostly from the “Fortune 500” list are collected to build the English profile corpus of American companies (hereinafter referred to as the EPCACs). All the profiles are downloaded from the official websites of these companies, saved in TXT form and encoded using the UCS Transformation Format-8-bit (UTF-8).

Table 1. List of chinese companies

Huawei	Inspur	Tongfang	Changhong	INESA	NARI Group
Lenovo	Xiaomi	SMIC	Hisense	ZTT	FiberHome
Haier	Unisplendour	Henan Senyua	BYD	Skyworth	Futong
TCL	Hengtong	Aisino	Founder	Tongding	O-film
ZTE	Hikvision	Jinglong	BOE	Tongfang	Datang

Table 2. List of american companies

Arrow	Rochester	Micro Focus	Charter Communications	Kingston	Tech Data
Qualcomm	CyberPowerPC	HP	Cisco	Kodak	AT&T
Monoprice	GE	Dec	Cognizant	Microsoft	IBM
Belkin	Honeywell	Dell	Comcast	Avnet	Intel
CenturyLink	Verizon	Sonos	Quantum	Motorola	HP

Research Questions

(1) What are the similarities and differences between the English profiles of Chinese companies and those of the American companies in word frequency, key-words, STTR, lexical density, keywords, average sentence length and passage length etc.?

(2) What are the implications of the differences in language features for the translation of the profiles of Chinese companies in the electronic and information industry?

Research Tools

AntConc 3.4.4, Claws 7, WordSmith 4.0 and other tools are used to compare the two corpora in terms of their language features including word frequency, key-words, lexical diversity and density,

STTR, mean sentence length and passage length, etc.

Findings

Word Frequency

Word frequency can be used to compare the frequency of words in different text files. By means of Antconc 3.4.4, the word lists of EPCCCs and EPCACs are set out in frequency order. It is found that of the 50 most frequent words of the two corpora, 27 are shared, including content words like *technology, products, services*, which suggest both American and Chinese companies attach importance to these aspects in their profiles. The difference is that words like *China, group, development, global, top, enterprise, innovation, smart, national, research, market, information, Chinese* appear significantly more often in the profiles of Chinese companies, while *customers, solutions, software, mobile, people, help, cloud, quality* appear more frequently in those of American companies. Besides, in the profiles of American companies, *we* and *our* are among the top 10 most frequent words.

Key-Words

Key-words provide a useful way to find out which words characterize the texts analyzed. In this study, the Brown Corpus is adopted as the reference corpus. It's found that the two corpora share 15 of the top 50 key-words, including *technology/technologies, innovations, product/products, company/companies, and, internet, digital, solutions, world, global, cloud, customers*, which suggest that both Chinese companies and American companies attach great importance to their products, customers and the global market. As the companies are all from the electronic and information industry and it's understandable that the frequency of these words such as *technology, innovations, internet, digital, solutions, cloud, etc.* is quite high. Apart from these similarities, the two corpora also show some differences in key-words. Of the top 50 key-words of the EPCCCs, 13 are company names such as *ztt, tcl, byd, etc.* while there are only 6 company names in those of EPCACs. Besides the company names, the words suggesting the location of the companies such as *China, Chinese, Shenzhen, Shanghai* are also among the top 50 key-words of the profiles of Chinese companies. In comparison, the words characterize the profiles of the American companies are *we, our, software, mobile, services, business, connect, challenges, partnerships, sustainable, etc.*

Lexical Diversity and Lexical Density

Lexical diversity is a measure of how many different words that are used in a text, while lexical density provides a measure of the proportion of lexical items (i.e. nouns, verbs, adjectives and some adverbs) in the text (Victoria Johansson, 2008: 61). Both measures are easily accessible and they provide important insights into the texts. Both lexical diversity and lexical density have been shown to be significantly higher in writing than in speaking (Ure 1971, Halliday 1985).

Lexical diversity is often used as an equivalent to lexical richness (Daller, Van Hout & Tredders-Daller 2003). The traditional lexical diversity measure is type-token ratio (TTR), the ratio of different words (types) to the total number of words (tokens) (Lieven 1978, Bates, Brettierton & Snyder 1988). But TTR varies widely in accordance with the length of the text or corpus of texts which is being studied. Therefore, a different strategy for computing this was adopted, i.e., the standardized type/token ratio (STTR), by means of WordSmith 4.0. In this study, the standardized TTR of the profiles of American companies is 45.67, slightly higher than that of the Chinese Companies, 43.94.

According to Ure (1971: 443-452), lexical density constitutes the estimated measure of content (nouns, adjectives, verbs, adverbs) per functional and lexical units in total. Ure concludes that a large majority of the spoken texts have a lexical density of under 40%, while a large majority of the written texts have a lexical density of 40 % or higher. The higher the lexical density, the more information the text contains. Spoken texts tend to have a lower lexical density than written ones, and texts with a higher lexical density are usually more difficult to comprehend. By means of

AntConc 3.4, the lexical density of the two corpora is calculated as follows (see Table 3).

Table 3. Comparison of lexical density between EPCCCs and EPCACs

	Nouns	Lexical verbs	Adjectives	Adverbs	Lexical word tokens	All tokens	Lexical density
EPCCCs	4390	942	1048	178	6658	11767	56.58
EPCACs	1168	387	326	58	1939	3373	57.49

In terms of the measures of lexical diversity and lexical density, the two corpora do not show much difference with the results of the EPCACs slightly higher than those of the EPCCCs, which suggests that the English profiles of Chinese Companies and those of the American Companies are similar in lexical richness and in the amount of information contained.

Mean Sentence Length and Passage Length

Butler (1985: 121) proposed that sentences are grouped into three categories by length: short (1-9 words), medium (10-25 words) and long (more than 25 words). Generally, the longer the sentence is, the more difficult it will be comprehended and the more formal and precise it will be.

Table 4. Comparison of mean sentence and passage length between EPCCCs and EPCACs

	Total number of words	Number of sentences	Mean sentence length	Mean passage length
EPCCCs	11767	484	24	392
EPCACs	3373	186	18	112

From table 4, it can be found that both the mean sentence length and mean passage length of the profiles of Chinese companies exceed by far those of the American companies.

Discussion

Through the comparison of the word frequency and the key-words of the two corpora by means of AntConc, it is found that the English profiles of Chinese companies and those of American companies have both similarities and differences in the words characterizing the texts. They all attach great importance to their technology, innovations, products, internet, solutions, customers etc. in their profiles. Besides, in the profiles of American companies, “we” and “our” rank both the top 10 most frequent words and the top 10 key-words, which suggests that they mainly adopt the first-person perspective, for example:

We help customers use technology to slash the time it takes to turn ideas into value (HP).

You may know us for our processors. But we do so much more (Intel).

By contrast, the Chinese companies use their company names more frequently. Comparatively speaking, when using the first-person perspective, the profiles sound more personal while when adopting the third-person perspective, the profiles sound more formal. Substituting “we” for the company name in the profiles can make an emotional appeal that helps forge a relationship between the company and the readers of the profile. It can also be found that the profiles of American companies are more customer-oriented as the key-words show that they focus more on customers, solutions, services, quality, reliability and partnership, etc. for example:

We are constantly innovating, looking for better ways to serve our customers (AT&T).

We help customers maximize the value of these assets to achieve their goals (Quantum).

A closer study of the language of the profiles reveals that the American companies have made the best of the customer-centric approach with very friendly language style.

In terms of lexical diversity and density, the two corpora do not show much difference, which suggests that the English profiles of Chinese Companies and those of the American Companies are similar in lexical richness and in the amount of information contained. The two corpora both have a relatively high lexical density based on Ure’s conclusion about the lexical density of spoken and

written texts (1971: 443-452). The reason is that company profiles often contain a lot of information such as the history, the products, the operations and future goals of a company.

The mean sentence length and passage length of the two corpora indicate that English profiles of American companies are much more concise and readable than those of the Chinese companies. One of the reasons is that most of the American companies are already world famous and that the reputations of their products have been established, so they can focus more on how to meet the needs of the current and potential clients. Their history, achievements and honors are often put in other subcategories rather than in the profiles. But for the Chinese companies, they are well known in China but not so worldwide. Thus it is necessary for them to talk more about their development, achievements, awards and honors to assure the foreign customers that their company can be trusted and their products are of high quality. However, this necessity does not justify the redundant information in their English profiles.

Conclusion and Implications

Through the comparative analysis, it can be concluded that compared with the English profiles of Chinese companies, those of the U.S. companies embody stronger target audience awareness with more concise and readable texts in friendlier language style. These findings shed light on how to improve the English profiles of Chinese companies:

The Principle of Conciseness.

Both the mean sentence length and passage length of the profiles of American companies are considerably shorter than those of the Chinese companies. Moreover, a close look at the English profiles of Chinese companies reveals that they are mostly translated from the Chinese versions sentence by sentence and contain a lot of redundant information, ignoring the communicative functions of the translation and the needs of the target audience. Therefore, it is recommended that the translators refer to the English profiles of famous American companies in the same industry as the paradigm, delete the redundant information, and if necessary, rewrite the profile in English in order to improve the readability of the text.

The Principle of Customer Orientation.

The English profiles of Chinese companies focus too much on their history, achievements and honors, while those of the American companies are more customer-oriented. Therefore, it is strongly recommended that the Chinese companies should focus more on how they can help them by providing them with quality products and good services.

Use the First-Person Perspective When Necessary.

Though the repeated use of the company name in the profiles may help enhance the readers' impression of the company, it also results in the content being less accessible and comprehensible. In this study, the profiles of American companies serve as a good example of using the first-person perspective to reinforce comprehension and establish a personal connection with the readers.

Focus on the Target Market.

The key-words of the English profiles of Chinese companies indicate that they have put too much emphasis on the domestic market, ignoring its connection with the target market. As the target readers of their English profiles are the overseas clients and business partners, thus the priority should be given to the target market in order to make the contents of the profiles more relatable to the target customers.

Acknowledgement

This research is supported by the "Key Platform and Major Program of Higher Education Institutions in Guangdong Province (2016WQNCX193)".

References

- [1] Baker, Mona, “Corpus Linguistics and Translation Studies: Implications and Applications,” In M. Baker, G. Francis, E. Tognini-Bonelli (eds). *Text and Technology in Honor of John Sinclair*, Philadelphia/Amsterdam, John Benjamins, 1993, pp. 233-250.
- [2] Bates, E., I. Bretherton & L. Snyder, *From first words to grammar: individual differences and dissociable mechanisms*, Cambridge: Cambridge University Press, 1988.
- [3] Daller, Helmut, Roeland van Hout & Jeanine Treffers-Daller. “Lexical Richness in the Spontaneous Speech of Bilinguals”. *Applied Linguistics* 24(2), 2003, pp. 197-222.
- [4] Huang Youyi, “Adhering to the Principle of Three Closeness and Handling Difficulties in Translating Overseas-Targeted Publicity Materials,” *Chinese Translators Journal*, 2014, pp. 27-28.
- [5] Johansson, Victoria, “Lexical diversity and lexical density in speech and writing: a developmental perspective”. *Working Papers (53)*, Department of Linguistics and Phonetics, Lund University, 2008, pp. 61-79
- [6] Laviosa. Sara, *Corpus-based translation studies: theory, findings, applications*. Amsterdam, Rodopi, 2002, pp.5.
- [7] Lieven, E. V. M. “Conversations between mothers and young children: individual differences and their possible implication for the study of child language learning”. In N. Waterson & C. E. Snow, *The development of communication*, Chichester: Wiley, 1978, pp.173-187.
- [8] Li Quandong, “Studies on translation of Chinese company profiles into English”. Shanghai: Shanghai International Studies University, 2013.
- [9] Stubbs, M. “British Tradition in Text Analysis: From Firth to Sinclair”. In M. Baker, G. Francis, E. Tognini-Bonelli (eds). *Text and Technology in Honor of John Sinclair*, Philadelphia/Amsterdam: John Benjamins, 1993, pp.8-13.
- [10]Ure, J. “Lexical density and register differentiation”. In G. Perren and J.L.M. Trim (eds). *Applications of Linguistics*, London: Cambridge University Press, 1971, pp. 443-452.
- [11]Wolfgang Iser, *The Act of Reading: A Theory of Aesthetic Response*. Baltimore and London: Johns Hopkins University Press, .1978.
- [12]The List of the“2017 China’s Top 100 Enterprises in Electronic and Information Industry”. Available at: https://www.guancha.cn/TMT/2017_08_01_420947.shtml