

中国英译旅游宣传材料的语料库分析

摘要

旅游业在全球范围内的迅猛发展使不同文化背景的人们进行文化交流变成了一种需要。在我国，旅游英语特别是网上的旅游英译材料是外国游客接触和了解我国的媒介。因此，旅游材料的翻译质量举足轻重。而目前旅游文本的翻译研究在中国还是一个比较新的话题，近年来主要是结合最新语言学研究的成果来对旅游资料进行跨文化对比分析或题材分析，但大多数只是翻译实践的总结，缺乏系统理论的指导。

本论文试图对互联网上的中国旅游翻译文本和英美国家的旅游文本进行对比，找出它们的不同，以探讨中国旅游材料英语翻译的有效途径，促进中外交流。为了使研究更科学，更系统化，作者采用了语料库语言学的方法。

语料库为研究者提供了大量真实的语言数据库从而使得语言分析研究能够基于自然发生的语言结构和形式而不是靠自觉感知的语言现象进行。语料库语言学不是语言学的分支，而是一种崭新的语言学研究方法。研究者使用电脑对大量文本进行分析处理，从而对语言和语言的使用得出结论，这单靠对某些文本的分析和某些语言学者的直觉是无法做到的。

论文共分六章。

第一章：导言，介绍研究的背景、目前国内旅游翻译文本存在的问题，研究目的和研究方法。

第二章：对旅游材料作简要介绍，包括旅游材料的定义、不同的分类方法、主要功能以及语言特点，并结合具体的旅游材料进行了论述。

第三章：回顾了国内外旅游材料翻译的发展历程。其中国内的学者主要从文化、功能、美学、翻译实践的角度进行了研究和总结。作者对他们的探索做出了客观的评价，并指出了存在的不足。

第四章：系统介绍了语料库语言学的历史渊源和发展沿革，语料库语言学的特点及广泛应用。此外，作者还简要说明了语料库研究中涉及到的技术性问题，如语料库的建立，计算机软件，统计工具等，为下一章的具体分析奠定了基础。

第五章：对中西旅游材料的语料库分析。首先，语料库的建立，包括抽样的标准、样本的数量和比例、样本的有效性。然后，利用 awk 软件对两个语料库的几种文本数据，如词汇量/总词数比、句长进行了计算和分析。通过比较，作者发现总体上中国旅游翻译文本的词汇量/总词数比（Type-token ratio）和英美国家的旅游材料相差不大，甚至在某一类文本上还高于后者。但是个案分析显示中国旅游材料中专有名词使用过多，在一定程度上造成了阅读障碍，影响了文化的交流。在平均句长方面，中国旅游材料略高于英美旅游文本，但是句型明显缺少变化，给人一种呆板的印象。此外，作为对话料库研究的补充，作者还就中西旅游材料的风格差异进行了探讨。

第六章：根据前一章的分析结果，针对词汇量/总词数比、句长和风格三个方面作者提出了汉语旅游文本翻译的建议。

关键词：中国旅游英译材料；英美旅游材料；语料库语言学

Translation of Chinese Tourist Publications:

A Corpus-focused Study

Abstract

This thesis makes a contrastive study between the translated versions of Chinese tourist publications and the English source texts, attempting to explore the differences between them and then proposing effective strategies to advance the translation of Chinese tourism culture. In order to make the research more scientific, the author employs corpus linguistics as the main theoretical framework.

A corpus provides a large empirical database of discourse so that analyses are based on naturally occurring structures and patterns of language use rather than intuitions or perceptions. Rather than a linguistic discipline, corpus linguistics is a brand-new methodology for linguistic research. With the use of computers, researchers are allowed to work with a large variety of texts and thus to seek generalizations about language and language use which can go beyond particular texts or the intuitions of particular linguists.

The paper, divided into six parts, deals with the following issues:

Chapter One presents the research background, the research problems, the research objectives based on the research problems and the methodology employed in the dissertation.

In Chapter Two, some basic knowledge about tourist publications is discussed, including their definition, classification, basic functions and language features.

Chapter Three takes a historical look at the previous research concerning tourist publications translation.

In Chapter Four, corpus linguistics, employed as the theoretical foundation of the present research, is analyzed in detail. The detailed analysis includes definitions, analysis tools and statistical measures, which serve as the key methods of the research.

Chapter Five is contributed to the corpus-based analysis of tourist publications on

the Internet. First, two mini-corpora are constructed, consisting of comparable texts (tourist texts in English translated from Chinese and English source texts) With the assistance of statistical package and computer programs, two aspects (type-token ratio and sentence length) are investigated. In addition, stylistic features of each are explored through a case study.

In Chapter Six, based on the studies in the previous chapters, is about how to make Chinese tourist publications more readable and more appealing to foreigners. Strategies are proposed in such respects as diction, sentence length and style.

Key words: the English version of Chinese tourist publications, English tourist texts, corpus linguistics

Acknowledgement

I owe great gratitude to my supervisor, Prof. Deng Hongfeng, who has given me so much constructive instruction and insightful guidance throughout my research. In the course of writing the thesis, the frequent discussions with him have encouraged, supported and enlightened me. I am indebted to him not only for his giving me invaluable guidance, and warm-hearted encouragement, but also for his arousing my interests in the field of translation of tourism texts with culture, as well as his providing me with his profound knowledge and precious materials in this field during the period of my study at Ocean University of China.

My gratitude also goes to Miss Gao Linghui, Miss Zhao Chunxia, Miss Wang Guiling and all my friends who have offered me a lot of reference materials.

Moreover, the present thesis has benefited from a large body of related works as seen in the bibliographies.

Last but not least, I would like to thank my parents and my husband for their love and support.

Chapter 1 Introduction

1.1 Research Background

In my postgraduate study, I have access to English versions of China tourist publications. Compared to the breathtaking scenic pictures in them, the language is frustrating. Later on, with my supervisor's proposals and guidance, I began to collect materials of the translated version of China tourism culture and materials of English-speaking countries. Comparatively, there exist striking differences between them in some linguistic parameters. In addition, I find the English translation leaves much to be desirable and the publicity function of these materials cannot be effectively fulfilled. There exist many problems, such as linguistic errors, pragmatic errors and cultural translation errors, which often prevent one from thoroughly understanding the text.

All these problems can't be ignored, for tourism in today's China is soaring owing to the rise of multinational exchanges and corporations. According to the data reported by China National Tourism Administration (CNTA), the number of tourists entering China in 1990 reached 72.8 million, of which 8.43 million were foreign tourists, 40 times the figure for 1978. Currently, China has become an important tourism destination in Asia, and the fifth large tourism country in the world. With its rapid development, more and more foreigners are becoming interested in China's long-standing history and unique culture, beautiful scenery and attractive historical sites. Under this circumstance, as the principal means of publicizing tourist attractions and promoting tourism in China, tourist publications play a significant role and are needed to be translated into English in much larger quantity and higher quality.

An overall study on this topic, therefore, becomes an urgent necessity. As a matter of fact, the study of tourist publications translation remains new worldwide. In China, the translation of publications did not catch Chinese scholars' attention until early 1990s. And till now there has never been a single book wholly devoted to this field as far as the author knows. Considering the inadequate research on this topic, the author intends to make a tentative study on the differences between the English versions of China tourist

materials and English source materials under the principles of corpus linguistics, which can serve this end well.

1.2 Problems and Purposes of the Research

Despite the importance of tourist materials translation, the present state of translation is far from satisfactory. Serious mistakes may frequently occur in some publications, which results in a negative influence on the image of our country. During these years, some translators have written articles calling on much attention on the translation of Chinese publicity materials including tourist ones. They claimed that poor translation had led to communication breakdown and eventually damaged the international image of China. Therefore, it is of great significance to make an in-depth research on tourist publications translation, probing into the causes of the poor translation and managing to improve the quality of the English version.

Duan Liancheng (段连城) in *Appealing for More Attention on External Publicity* proposes that the problems available in the translated version of publicity materials can be summarized as two “diseases”. The “symptoms” of the first kind of “disease” include numerous spelling errors, grammatical mistakes and inappropriate diction. On the other hand, the second kind of “disease” is characterized by the difficulty for foreigners to understand and comprehend tourist information texts, not to say to appreciate and enjoy them, due to striking cultural barriers. (段连城, 1992)

Among these problems, the first kind of “disease” is caused mainly by translators’ carelessness and poor basic language skills. Such mistakes are easy to be recognized and corrected and will not bring much trouble to the comprehension of the tourist texts, though they do make negative influence on the reader’s mind. Meanwhile, the causes of the second kind are much more complicated. In the author’s opinion, the main reason lies in the fact that the translator has just rendered the materials literally, without considering the cultural elements of the texts. Bearing no cultural consciousness in mind, the translator pays but little attention to the unique Chinese cultural background and neglects the tremendous cultural differences between the source text reader and the target text reader, who belong to different cultural groups. It is the cultural factors instead of

linguistic ones that lead to the failure of such kind of translation. More important, as the problems are caused by cultural factors, they are not easy for native translators or readers to detect and correct, only resulting in the fact that foreign readers are lost in a sea of Chinglish and weird descriptions

The thesis is intended to achieve three goals: First, probe into some linguistic and stylistic features of the West and China in tourism culture based on a large database. Second, compare and analyze the differences between them. Last, suggest various strategies to bridge the gap based on the two steps above.

1.3 Methodology for the Research

Methods are always dictated by purposes. Therefore, corpus linguistics, a brand-new but increasingly popular approach plays a dominant role in this research. Corpus is a structured collection of texts specially compiled for linguistic analysis. (Kennedy, 2000:3) Corpora have been compiled for many different purposes, which in turn influence the design, size and nature of the individual corpus. Generally, they are designed for linguistic research on prosody, lexis, grammar, discourse patterns or pragmatics.

“Corpus linguistics is based on bodies of text as the domain of study and as the source of evidence for linguistic description and argumentation.” (Kennedy, 2000:7) Rather than a linguistic paradigm, corpus linguistics is a way of doing linguistics, "a methodological basis for pursuing linguistic research" (Leech, 1992). Corpus linguistics did not begin with the development of computers but there is no doubt that computers have given corpus linguistics a huge boost by reducing much of the drudgery of text-based linguistic description and vastly increasing the size of databases used for analysis. (Kennedy, 2000) Computers have permitted linguists to work with a large variety of texts and thus to seek generalizations about language and language use which can go beyond particular texts or the intuitions of particular linguists. It can apply to literary studies, lexicography, grammar studies, translation, language education, etc. (ibid)

To conclude, corpus linguistics has become a mainstream. That's why the author tends to study the tourist publications translation based on the corpus-based analysis.

Beginning with the construction of two corpora composed of "comparable texts" (tourist texts in English translated from Chinese and the corresponding English source texts), a contrastive research on the lexical, sentential and textual features of each will be carried out. Then the reasons for those differences will be discussed, followed by tentative guidelines for tourist publications translation.

1.4 Thesis Structure

This thesis comprises six chapters.

Chapter One presents the research background, the research problems, the research objectives based on the research problems and the methodology employed in the dissertation.

In Chapter Two, some basic knowledge about tourist publications is discussed, including their definition, classification, basic functions and language features.

Chapter Three takes a historical look at the previous research concerning tourist publications translation.

In Chapter Four, corpus linguistics, employed as the theoretical foundation of the present research, is analyzed in detail. The detailed analysis includes definitions, analysis tools and statistical measures, which serve as the key methods of the research.

Chapter Five is contributed to the corpus-based analysis of tourist publications. First, the process of corpus compilation in this research is illustrated, including the criteria, proportion and size of sampling. In addition, the computer program and the statistical software used are introduced in detail. Based on the two corpora, type-token ratios and sentence length are analyzed and compared. Apart from that, stylistic features of each are explored through a case study.

In Chapter Six, based on the studies in the previous chapters, is about how to make Chinese tourist publications more readable and more appealing to foreigners. Strategies are proposed in such respects as diction, sentence length and style.

Chapter 2 Introduction to Tourist Publications

2.1 Definition of Tourist Publications

According to the World Tourism Organization (WTO), “*Tourism comprises the activities of persons traveling to and staying in places outside their usual environment for not more than one consecutive year for leisure, business and other purposes*” (向萍, 1998). It can be classified into two categories: material tourism culture and non-material tourism culture. The former covers a variety of things about tourism ranging from scenic spots, natural resources, climate, geographical locations, historic relics and figures, to food, transportation, facilities, etc. The latter contains customs, rites, attitudes and services, etc. Meanwhile, the tourist is a role taken by people of all social classes, and many of them do not share similarities culturally. Through traveling, tourists experience unique and exciting life in different cultural groups. Publication means “*the dissemination of information of promotional materials or the information with news value issued as a means of gaining public attention or support*” (Oxford Advanced Learners Dictionary of Current English, 1997:1167). In this thesis it is broadly a synonym of promotion. Thus, tourist publications supply basic and sufficient information on tourist destinations with an aim of publicizing and promoting the tourism products and they may appear in various forms including brochures, pamphlets, guidebooks, picture albums, tour maps, etc.

Like other means of communication, tourism publications have their own initiators and recipients. They offer information needed by potential tourists and investors and serve as a tool of their initiators to achieve their promotional ends. When people plan to travel, they usually resort to the tourist information distributed by the agencies concerned. The information provided in the publications usually involves the following major aspects: the attractions (the scenery, the local food, climate, and customs, history and culture), facilities, transportations and services as well. As tourism industry will not keep rapidly growing without sufficient tourist promotion, the promotional activities are quite useful and serve the function of informing, persuading and influencing the customer’s

purchase decision. Nowadays, with the increasing international communications, there has been found to be an urgent demand both in expanding the quantity and improving the quality of the Chinese tourist publications. It is, therefore, the translator's great task to render Chinese tourist materials into refined English version, With the purpose of promoting China's tourism worldwide.

2.2 Classifications of Tourist Publications

In terms of objective attributes, tourist resources can be classified into two categories: material tourism culture and non-material tourism culture. The former covers a variety of things about tourism ranging from scenic spots, natural resources, climate, geographical locations, historic relics and figures, to food, transportation, facilities, etc. The latter contains customs, rites, attitudes and services, etc.

American scholar Goeldner claims that tourist resources consist of five categories: cultural attractions, natural attractions, events, recreation and entertainment attractions (Goeldner, Ritchie and McIntosh, 2000) The detailed information is illustrated as the following figure:

Categories	Contents
Cultural attractions	historic sites, architecture, cuisine, relics museums, auditoriums, theatres, etc.
Natural attractions	mountains, parks, botanical gardens,zoos, river, lakes, beaches, islands, etc.
Events	festivals, community events, sporting events, religious events, etc.
Recreation	golf tour, traveling on foot, cycling tour, skiing tour, swimming, etc.
Entertainment attractions	theme park, amusement park, cinema, shopping facilities, etc.

Figure 2.1 Goeldner's classification of tourist resources

Thus, tourist publications are sorted accordingly. This classification is adopted as the main criterion of sampling in the corpus-based research.

In terms of the source, they include brochures from publishers, tourist materials distributed by agencies and bureaus and tourist information from the Internet. Samples used in the thesis mainly come out of the last source, since the Internet takes the priority of unparalleled accessibility.

According to the content, tourist publications can be broadly classified into three categories, i.e. brochures or pamphlets of tourist attraction, guidebooks of cities or countries, and information on tourist operations. In this thesis most samples belong to the first category, as they are the most typical tourist information texts among those categories.

2.3 Basic Functions of Tourist Publications

Tourist publications are addressed to a selected group of potential customers in order to inform them of the tourist products and services and persuade them to buy. Through these texts, authors intend not only to capture the prospective customers' attention but also to sustain their interest, and eventually to convince them of the benefits of the product or service being promoted. In accordance with the above purposes, tourist publications are usually informative, attractive and persuasive in function.

2.3.1 Informative Function

“Language serves an informative function when it is used to tell what the speaker believes, to give information about facts, or to reason things out.” (胡壮麟, 1998)
Tourists need to acquire all kinds of information when they start a tour to a new place. Such information includes: geographical conditions, people, places of interest, culture, history and customs, etc, which serves to arouse the reader's interest and result in an impulse toward the final traveling action. Thus a tourist information text has to be informative to achieve its final persuasive function.

Sample 2.1

Cave and Basin National Historic Site of Canada

The Cave and Basin National Historic Site commemorates the birthplace of Canada's National Park system, which began here in 1885 when, to settle an argument over the site's ownership, the government declared that the springs would belong to all Canadians -- as part of Canada's first national park.

Attractions

The main attractions at Cave and Basin are the naturally occurring, warm mineral springs that can be found inside the cave, and outside in an emerald coloured basin. The growth and development of Canada's national parks are celebrated in interactive displays and exhibits on site. Visit the replica of the 1887 bathhouse to see the "Steam, Schemes and National Dreams" film on the discovery of the site and the establishment of Canada's national parks. There are self-guiding "Discovery" and "Marsh" interpretive boardwalk trails where you can discover the rich history of the hot springs and explore the flora and fauna, and a paved cycle path to Sundance Canyon. The site is open from October 1 through May 14.

For more information:

Cave & Basin Historic Site

P.O. Box 900

Banff, Alberta Canada T1L 1K2

Phone: (403) 762-1566

Fax: (403) 762-1565

Email: banff.vrc@pc.gc.ca

----From <http://www.pc.gc.ca>

In this passage, the detailed information of the tourist destination— *Cave and Basin National Historic Site of Canada* is available to tourists, e.g. the main attractions, address, telephone number, website, email and facilities.

2.3.2 Attractive Function

Attractive function means tourist publications should be capable of producing sensations and then drawing the consumer's attention. There is no doubt that a tourist information text cannot attract people and stimulate purchasing merely with the help of its informative function. A good text must not only provide information about tourist attractions, but also please people. Such pleasure helps people memorize the content of the text and become willing to visit the specific place, i.e. to buy the tourism product.

Well-written publications should be lively and vivid so that they can present the most outstanding features of the tour and can be easy for prospective tourists to remember. Generally speaking, to realize its attractive function, a good tourist information text should be of aesthetic value and memory value.

Sample 2.2

This winter tour to three provinces in northeast China is unique and interesting. Pastures, mountains and trees are covered by snow, a sea of whiteness. Local ice sculptures are another attraction. Skating, skiing, hunting and horse-back riding offer special enjoyment. After a day's activity, free chats with the local people on warm earthen beds help rid travelers of the cold.

-----From <http://www.cnta.gov.cn>

No doubt such a thrilling description will be a great stimulus for every reader to take the tour even right now.

2.3.3 Persuasive Function

The ultimate purpose of the tourist information text is to motivate and persuade customers to buy the tourist product; that is, to visit the place. The producer of the text intends to stimulate the reader's interest, turn them into prospective tourists and induce them to act. Thus, persuasive function should be a fundamental one for tourist information texts. However, to realize the persuasive function, the text must first perform its informative and attractive functions. In other words, only when the texts are informative and attractive can they interest people and persuade them to take action. The above three functions of tourist publicity materials are closely interrelated and are all of great importance to fulfill the ultimate purpose.

2.4 The Typical Language Features of Tourist Publications

Tourist publications are characterized by their free and mixed style which may be descriptive, narrative, emotive and even poetic. Most tourist materials are descriptive or narrative. The language in these tourist information texts is appropriate written English, not too colloquial or too formal, and can be well understood by tourists of different classes and with different educational backgrounds. The common features of tourism

English can be summarized as follows: (1) brief, (2) vivid and appealing, (3) humorous, (4) prominent use of adjectives and adverbs, (5) figurative use of language.

a. Brief: A brief style is a must-have property of tourist publications. Phrases, short sentences are more frequently used than clauses since they better convey the main points of the tour and make it easier for the tourists or potential tourists to remember.

Sample 2.3

Surfboards and boogie boards decorate this wacky shack. The food's great - and certainly no wipe out!

- Great pizza – like the Big Kahuna Hawaiian, Awesome Pepperoni and Wipe Out Cheese
- Salad, beverages, and more
- Sit outside and catch some rays!

Cuisine: Vegetarian, Healthy Selections, Italian

Service Type: Quick Service

Price Range: \$ (\$10 per person or less) *

Meals Served: Lunch, Dinner

Location: Paradise Pier

----Form [http:// www.usa.worldweb.com](http://www.usa.worldweb.com)

b. Vivid and appealing: Only when a tourist text reads vivid and appealing can it impress the reader most. To draw the tourist's interest, the writer usually uses the words and expressions that are specific and vivid as they are more appealing than general terms. They will bring readers a deeper impression by depicting vivid pictures in their mind as shown below:

Sample 2.4

Soarin' Over California

Hold your breath as you experience the ultimate sensation of free flight. Your soarin' journey takes you from one end of the Golden State to the other in this awe-inspiring scenic tour.

- Take in immersive breath-taking aerial vistas while flying in this simulated hang glider
- Your toes dangle just above the awe-inspiring landscapes such as the Golden Gate Bridge and Yosemite
- Feel the wind in your hair as you see California as you've never seen it before!

Category: Big Thrills

Location: **Golden State**

Height Requirement: **40" (102 cm) or taller**

This attraction has **Disney's FASTPASS® service**

-----Form [http:// www.usa.worldweb.com](http://www.usa.worldweb.com)

c. Humorous: In the texts of tourism culture, humor is the best-loved ingredient to tourists.

Sample 2.5

It is the only way to get the "feel" of our immense land. Just do not be surprised a bit of your heart remains behind when you depart.

d. Prominent use of adjectives and adverbs

Most prominent uses in tourist publications are the use of a large number of adjectives and adverbs. "Not all tourism culture are illustrated with color photographs, so you must use your vocabulary to convey impressions of the subject through precise adverbs and adjective"(Macdonald, 2002:301). The tourists' eyes, ears, mind and soul are often bombarded with adjectives and adverbs such as vast, beautiful, wonderful, fantastically, incredibly, greatly, etc. Pick up a random handful of brochures at any travel agent and you will find they are full of picturesque villages, golden sand, sun-drenched islands, and delicious cuisine. Therefore sometimes, overuse of adjectives and adverbs cause doubts—or even irritation.

Sample 2.6

Guilin's Landscape

The world-renowned scenic city of Guilin is naturally a fine tourist city in China. A distinctive karst landmass is the reason behind Guilin's fabulous landscape that is characterized by green mountains, sparkling waters, strange caves and statuesque monoliths. Xing'an County, 60 km to the northeast of Guilin, is the site of the Lingqu (a world-famous ancient canal), and the Leman Geological Scenic Zone where visitors go for both recreation and leisure. Ziyuan County, 149 km north of Guilin, features a landscape corridor where the scenery is by turns mysterious, breathtaking, strange and graceful. Longsheng County, 97 km from Guilin, features the folkways of the ethnic Zhuangs and Yaos, hot spring sanatoriums, and the world's most spectacular Dragon-Ridge Terraced Fields.

-----From <http://www.cnta.gov.cn>

Sample 2.7

A Unique Tropical Nature Preserve and Sanctuary

Hawaii Tropical Botanical Garden is a spectacular Garden in a Valley on the Ocean - acclaimed as one of the most beautiful areas in all Hawaii. The Garden is located on the Big Island of Hawaii, 8 1/2 miles north of Hilo on the four-mile Scenic Route at Onomea Bay. In this garden valley, nature trails meander through a true tropical rainforest, crossing bubbling streams, passing several beautiful waterfalls and the exciting ocean vistas along the rugged Pacific coast.

The Garden displays a vast variety of palms, heliconias, gingers, bromeliads, and hundreds of other rare and exotic plants from all parts of the tropical world - presently more than 2,000 species, and the collection is always growing! This non-profit nature preserve is dedicated to providing a plant sanctuary, a living seed bank, and a study center for trees and plants of the tropical world and to preserving the incredibly beautiful natural environment of Onomea Bay for generations to come.

----Form [http:// www.usa.worldweb.com](http://www.usa.worldweb.com)

e. Figurative use of language

Effective language is often closely connected with rhetoric. For example, metaphor and simile create images in readers' minds by comparing things that are usually not associated. Metaphor describes something in terms of something else, and they imply comparisons. Metonymy means "change of name," that is, the use of one thing for another related to it. Similes are more direct comparisons commonly identified by use of like or as. This occurs most often by using some sort of reference commonly known by tourists. Personification means to give life to the inanimate or sympathy to brute creation (陈淑华, 1990:117). For most tourists, such writing is a welcome diversion from routines, which might help readers remain interesting in reading.

Sample 2.8

The spring of the West Lake is very poetic, especially the famous scenery of the White Causeway in Early Spring. Built in Tang Dynasty, more than 1,000 years ago, the White Causeway is covered with silk-like willows and pink peach blossoms in spring, just like a bright and beautiful strap across the lake. Seeing such beautiful scenery, visitors would feel as if they have entered a fairyland on earth. (Simile)

Malaysia—A Shoppers' Heaven (Metaphor)

This year, Shopping Carnival promises even more as it goes nationwide from October 3—November 1. (Personification)

With more jewelry shops per square kilometer than any other city, Hong Kong is lined in gold, diamonds, pearls and precious gems. (Hyperbole)

Guilin's beauty is beyond praise. (Hyperbole)

The beauty of Guilin is more than I can describe. (Understatement)

Life is a journey. (Pun)

We can see the babbling water in the brook, hear the frogs croaking, cuckoo cuckooing, birds singing, sparrows chirping in the forest. All are lovely. (Onomatopoeia)

In parts of the sparsely settled western deserts and mountains, the minorities are the majority. (Alliteration)

Taste the food of emperors

Listen to the children singing

Smell the temple incense (Parallelism)

Studying the language features of English tourist publications certainly sheds light on translating Chinese tourist materials into English in a cross-cultural perspective. It paves the way for translating them in a new angle, i.e. taking intercultural communication factors into account. As we have studied the salient features of English tourism publications, in our translation, we should try to adapt the Chinese language style to the English one, in order to make it possible for foreign tourists to understand and appreciate easily.

Chapter 3 Literature Review

3.1 An Overview

In the past two or three decades, many scholars and experts have contributed to tourist publications translation. They apply a variety of theories to this new research, such as functional grammar, cultranslation, semiotics, semantics, psychology and discourse analysis, etc, all of which facilitate to broaden the horizons of scholars and to bring fresh air into the concepts of language, culture and translation. All in all, more and more attention has been paid to this field and greater emphasis has been placed on “equal response” of the target language readers. Such new perspectives on the theoretical front as well as the fairly extensive developments in specific interlingual contrastive studies have promoted considerably the understanding and mastery of the nature and methods of tourist text translation. However, these theories are not sufficiently applied to the translation of Chinese tourist literature, which, consequently, has benefited so little from recent developments in linguistics. But in the following literature review, we can still find the efforts of some experts who have attempted to use modernist linguistic theories to explore the translation of tourist publications.

3.2 Researches Abroad on Tourist Publications Translation

The study on tourist literature abroad is a comparatively new topic. Two typical representatives are Rosa Lore’s Sanz and Erik Castello.

Rosa Lore’s Sanz wrote an article: the Translation of Tourist Literature: the Case of Connectors.

This article reports on a small-scale study, which focuses on the analysis of connectors (conjunctions and discourse markers) in thematic position in English translation of Spanish tourist literature. The concept of “theme” adopted here draws on the explicit descriptions provided by systemic grammar. The study is carried out on a corpus composed of “comparable texts” (texts written in English and texts in English translated from Spanish) and of the corresponding Spanish source texts. It is argued here that the analysis of patterns of theme selection provides important insights into the

writing and, therefore, translation of a specific genre. (Sanz, 2003)

In the book *Tourist Information Texts: a Corpus-based Study of Four Related Genres*, Erik Castello aims to furnish a description of the language used in this kind of texts and of their structure, following a systemic-functional approach and making use of corpus linguistics. (Castello, 2002)

First, some of the contributions concerning lexical density and grammatical intricacy as well as other linguistic features are reviewed in the perspective of their application to the quantitative corpus analysis. Second, a detailed classification of the distinctions made between grammatical and lexical words as well as an accurate description of the way the clause-complex analysis has been carried out are also furnished. Third, by means of statistical and computational techniques it is shown how the frequencies of the linguistic features vary across four related genres, or subgenres, of tourist-information texts corresponding to four channels of transmission (brochures, magazines, tourist guides, and web-sites), thus typifying them. In the study the concept of genre is used as a theoretical framework, accounting for all the findings emerging from the investigation and is able, therefore, to highlight variation as well as uniformity across the related text types of tourist publications identified. Lastly, a description of the context of situation in which tourist publications are produced and made use of is expounded. (ibid)

To sum up, the above-mentioned studies mainly adopt linguistic approaches such as functional grammar, genre analysis, etc. Linguistic perspective gives insights into this genre and enlightenment of the further study on the translation of this genre. Due to the lack of studies of Chinese unique language abroad, the literature review in this paper is mainly about Chinese scholars and their relevant writings. It is also recently that tourist literature has begun to be studied theoretically, which is far behind its development in practice.

3.3 Chinese Studies on Tourist Publications Translation

Most of the translation studies on tourist publications in China are related to cultranslation, functional analysis or even just the summarization of experiences in translation practices. A detailed introduction is as follows:

3.3.1 Culture-based Study of English and Chinese Tourist publications

Cultural connotation and cross-cultural differences are stressed in Fu Wen' s "A Brief Discussion on Cultural Connotation and Traveling Translation" (傅文, 2001) Bu Shaoxian in "History and Culture of Minority Vs Tourism Translation for Publicity" points out the historical and cultural elements in C-E translation of minority tourism publicity. Political aspect is also mentioned. He also provides some methods of translation based on his own experience and thought. (卜绍先, 2000) Xing Yi' s "On the Criteria of Translating Tourism Publications and Some Suggestions for Bettering Translation Quality" analyzes some problems in the present C-E translation situation in China and gives suggestions on the standardization of our translation. The author also discusses the requirements on how to do the translation of Tourist Literature well. (邢怡, 2001) Chen Hong in "Analysis of Techniques in Tourist Material Translation" gives us feasible methods of translation, which however appear to lack theoretical framework to support. (陈红, 2003) Jia Shunhou, in his article "On the Cultural Adjustment of the Translation of the Chinese Tourist Material", through the comparison between the nature of the tourist material and that of the translation, argues that translators should focus their attention on rendering the cultural message while translating, so that the translated versions may reflect Chinese cultural condition to excite the curiosity of the potential foreign tourists; Translation itself can be regarded as the approach and medium of cultural communication. Furthermore, the author also presents some related problems. Lastly, the author concludes one principle and two methods in translating tourist materials, and that the translation of proper nouns should be done according to the established usages, and meanwhile the author calls on an earlier establishment of a standardization research center for its translation. (贾顺厚, 2002)

Besides, the researchers approaching Chinese tourism texts translation from the perspective of culture include Pu Yuanming(蒲元明), Yang Shoukang(杨寿康) and Liu Huimei (刘慧梅). Pu mentions three ways of dealing with tourism translation: addition, deletion and rewriting. Regarding the method of deletion, he holds that "All these quotations, if not helpful in the target text, had better be deleted. The deletion will make

the translation more precise and to the point and easier to understand.” (蒲元明, 1987: 30-31). Liu and Yang put forward five methods: explanation, adding, analogy, deletion and paraphrasing. According to them, “The Chinese (especially classical Chinese), after narrating or describing, tend to quote from poems, well-known sayings and so on to prove their feeling and point of view. Such being the circumstance, it is necessary to make some deletions.” (杨寿康, 刘慧梅, 1996:10-14)

The dissertation “The English Translation of Chinese Tourism Texts from the Perspective of Cultranslation” written by Su Bing, explores the effective methods applied to the translation of tourism writing under the guide of the theory of cultranslation and find solutions to the problems existing in the present English versions of Chinese tourism writing in order to advance the communication and understanding of Chinese culture in the world. (苏冰, 2003)

3.3.2 Functional Analysis on Tourist Literature

He Shan, in her thesis for Master’s Degree, “The C-E Translation of Cultural Elements in Tourist Publicity Materials— A Functionalist Perspective”, first examines some basic knowledge about tourist publicity material including their classifications, basic functions. Then the author focuses on the cultural characteristics of tourist publicity materials on the basis of the relationship between language and culture. In addition, the translation difficulties caused by cultural discrepancies are analyzed in detail. According to the author, at lexical level the difficulties lie in culture blanks, or culture-loaded words and expressions. While at textual level, the problems are caused by language conventions due to different thought patterns and aesthetic values between Chinese and westerners. To solve these problems, relevant translation theories including functional equivalence, text typology and Skopos theory that are reviewed. Finally, it is concluded that Skopos theory supplemented by text typology theory can provide the present study with a most effective theoretical framework. In the light of the tourist-centered principle and the task of promoting Chinese culture, some practical strategies are put forward to bring the effect of the translated version into full play. (和珊, 2004)

Chen Daiqiu’s “Translation of the Cultural Load and Imagery in Chinese Tourism

Texts: A Functional Perspective” adopts similar approach. By resorting to Skopostheorie, the author probes into the translation of the cultural load and imagery in Chinese tourism texts. After an analysis of the tourism motivators of potential tourists, the author, concludes that one of the motivators of potential tourists is to experience new cultures. Therefore, the translator, in order to stimulate the potential tourists’ motivation to buy a targeted tourism product, should treat the cultural load or foreignness in such a way as to enable the target readers to see in the translation the cultural foreignness. By comparing domestication and foreignization, the author argues that the translator, in order to retain the cultural foreignness, should adopt foreignization, including literal translation and transliteration. Then, the author holds that, since translation is an “intercultural action” and thus cultural default or cultural vacuum may arise when foreignization is adopted, necessary compensations such as annotation, amplification and analogy should be made so that the target reader can bridge the cultural gap. Of course, according to Skopostheorie, the cultural load considered inadequate to the translation purpose of stimulating the target readers' interest should be deleted. (陈代球, 2005)

“Chinese-English Translation of Tourist Text in View of Function” written by Su Liqin adopts the functional approach of Peter Newmark. She upholds that tourist writing has vocative function and focuses on readers. Therefore, its English translation must have the same effect. In the process of translation, certain concrete methods like explanation, supplement, analogy, deletion and adjustment etc. can be adopted in order to achieve the functional equivalence between the original text and the corresponding English version. (苏丽琴, 2001) Ke Xixiang in “Skopos Theory and the Translation of Tourist Literature” adopts Skopos theory, and she gives abundant examples. But systematic theoretical analysis occupies too small proportion in her article. (柯细香, 2002) Guo Ruijuan, in her paper “Text Typology Theory and English Translation of Tourism Materials” employs the text typology theory which is advocated by German scholar K. Reiss and British scholar Peter Newmark, on the basis of the function of language. According to her, the translator should aim at fulfilling the communication of the vocative and informative functions of tourism materials. Her paper puts forward some principles and techniques in

translating such materials into English. (郭瑞娟, 2003) Liu Helin in his paper “The C-E Translation Art of Poems in Tourist Literature” expounds the translation of landscape poetry in Tourist Literature. (刘和林, 2003)

Concerned with those researches of Chinese scholars, their investigation and data collection are really a valuable try, which is different from some translators or scholars’ subjective illustrations, therefore their researches are more adequate and convincing.

3.3.3 Other Approaches of Studies

Other C-E translation studies on tourist literature are respectively concerned with different aspects and angles. Here’s the detailed.

1) On Tourism

Yin Xiaohong’s “Tour Guide Translation”(印晓红, 1997), Ji Junchao’s “On Tour Guide’s Translation and Interpretation” (纪俊超, 2002), are both on tour guide translation which are effective for the training and practical translating work of tour guides. Chen Baiying’s “Tourism English and Its Translation” deals with some features of tourism and some suggestions on its translation. (陈白颖, 2003)

2) On Translation Practice

Yao Baorong and Han Qi (姚宝荣, 韩琪, 1998) suggest five methods: paraphrasing, addition, analogy, deletion and creative translation. Xing Yi (邢怡, 2001) raises four criteria for the translation of tourism publications: naturalness (the translation should read natural), accuracy (the translator should, on one hand, avoid spelling mistakes, and on the other, make sure the words employed can convey correct information, aesthetic sense (aesthetic sense can be achieved by means of diction, sentencing, correct spelling and grammar and proper rhetoric devices) and cultural meaning (importance should be attached to the translation of cultural meaning). Lai Yu (赖余, 1986) forwards two suggestions: fully understanding the source text and making the translation easy to understand. Chen Gang (陈刚, 1987) thinks that the translator should gain a correct understanding of cultural load and take into account psychological factors of the target reader such as their nationality, sensitive words and so on. Later, Chen Gang (陈刚, 2002) puts forward some more specific methods. He holds that 1) *pinyin* should be standardized;

2) imagery, if any, embodied in landscape names should be conveyed in the translation; 3) when dealing with cultural load, the translator should always be oriented toward the dissemination of Chinese culture; 4) attention should be paid to pragmatic meaning; 5) emphasis should be laid on poetry translation; and 6) cultural load should be retained as much as possible.

Ye Miao in his article “The Translation of Names of Scenic Spots”(叶苗,1995), Jin Huikang “A Discussion on Guangdong Tourism Translation”(金惠康, 2003), Xiong Xiangyang in his paper “the Translation of Names of Ethnic Minorities Customs”(熊向阳,1998) Yi Zhou in his “Some Gains Acquired in the Practice of Tourism Translation”(一舟, 2002) Liao Hongzhong in his “A Study on the Translation of Jiangxi's Scenic Spots” (廖洪中, 2002) give us a large number of valuable suggestions on translation methods of names of scenic spots or ethnic minorities customs, which are very valuable to the standardization of the translators' practical translation work.

Their studies are abundant in practical translation examples and suggestions, but not systematic theoretical instructions, so they appear a little sort of empirical and subjective.

3) On Sino-Western Aesthetic Mode

Jia Wenbo(贾文波),Gao Jinling(高金岭)and Jin Hailin(靳海林)have made their studies from the perspective of landscape descriptions. Jia (贾文波, 2000) makes an investigation into the aesthetic differences between English and Chinese manifested in landscape descriptions. According to him, when translating Chinese landscape descriptions into English, the translator is expected to reduce subjective descriptions in Chinese to objective ones in English. Jin (靳海林, 1996) argues that a good piece of tourism translation should reproduce the beauty of feelings, forms and dynamic state and suggests ways to achieve the beauty in those aspects: a correct comprehension of the original for the reproduction of feelings; transformation of static state into dynamic state for the reproduction of dynamic state and sufficient attention to lines, colors, sounds and smells for the reproduction of forms. Gao Jinling in “The Translation of Descriptive Touring Materials in Light of the Sino-Western Aesthetic Modes”(高金岭, 2003) puts forward some views on how to translate suggestive descriptions. He upholds that description of scenic spots can be divided into two parts: expository and suggestive. In

his opinion, the former is quite easy and the latter is the difficult part in touring materials from the perspective of the differences of aesthetic modes between Chinese and westerners. These differences chiefly focus on the following concepts: representation vs. expression; poetic pictures vs. picturesque poetry; abstract concepts vs. concrete concepts; rich in images vs. exact in depiction; time-oriented description vs. space-oriented description. In conclusion, the author comes to conclude that the study of differences in aesthetic modes between the two languages will shed light on the translation of suggestive descriptions in tourist materials.

4) On Reader-Response

“The Agreement between Information and Acceptance”, written by Wen Jun, Deng Chun, Gu Tao and Jiang Yujia (文军, 邓春, 辜涛, 蒋宇佳, 2002) together, is based on the common problems in the present tourist translation. They investigate the responses and attitudes of foreign tourists from different countries. Then, through the data analysis, the various problems of the present tourist translation are shown out, the tourists’ psychological acceptable degree is investigated, and some valuable suggestions and objective basis are put forward to improve the tourist translation.

5) On genre Analysis

Wang Lixiu in her thesis for Master’s Degree, “A Genre Analysis of Tourist Information Texts: Implications for Teaching ESP Writing”, analyzes the brief tourist information texts in travel brochure as a sub genre of promotional genres with an aim to identify the rhetorical organization in the given genre, identifies the linguistic features chosen by the expert users of the genre to realize their communicative purpose and finally explains those choices. In particular, she tries to make out a workable and efficient model for the teaching of ESP writing. (王丽秀, 2002)

3.4 Problems and Shortcomings of Previous Researches

These researches are of big value for the further translation studies on both theory and practice since they give us such numerous translation practical experiences and authorities. Some of them even have applied the recent linguistic theories and methods to this discipline, which provide us with a lot more enlightenment. However, room still

exists for further and more incisive investigation. For instance, a more scientific approach needs to be taken into consideration.

3.5 Summary

In the chapter of literature review, the author summarizes the relevant researches both home and abroad on this relatively new genre—tourist publications from various perspectives. It can be seen obviously that the translation of Chinese tourist publications is much more complicated than the translation of other genres, since it deals with not only bilingual, but also bi-cultural and bi-social transference. (安春平, 2005) Therefore, it is generally accepted that translating Chinese tourist publications is rather a complex process subject to the influences of a great many variable factors. Both Chinese tourist publications and English tourist publications are not only written in different languages, but also represent different cultures. They differ greatly in terms of linguistic and social-cultural conventions. The translators shall not only pay their special attention to certain peculiar ways in which language generates aesthetic effects in Chinese and to the methods used to achieve expressive identities in English but also bridge the gap between the translated Chinese tourist publications and the corresponding English source texts. However, linguistic theory alone has failed to do the job intended. With the development of computational technology, nowadays the relatively better approach to studying the translation of Chinese tourist publications is the corpus-based approach. This paper intends to compare the major linguistic features between the translated version of Chinese tourist publications and the English ones based on corpus linguistics.

Chapter 4 Corpus Linguistics

4.1 Introduction

“Studies of language can be divided into two main areas: studies of structure and studies of use. Traditionally, linguistic analyses have emphasized structure—identifying the structural units and classes of a language(e.g., morphemes, words, phrases, grammatical classes) and describing how smaller units can be combined to form larger grammatical units(e.g., how words can be combined to form phrases, phrases can be combined to form clauses, etc.)” (Biber, 2000:1)

According to Leech (1992:107), however, “the focus of study is on performance rather than competence, and on observation of language in use leading to theory rather than vice versa.” Therefore, a different perspective is adopted in this dissertation, i.e., to emphasize language use. From this perspective, we can investigate how language is exploited respectively in English tourist materials and the translated version of Chinese ones. Rather than examining the differences between both theoretically, we study the actual language used in naturally occurring texts. This approach involves two central research goals: (1) finding typical language features of each style, and (2) comparing and analyzing the differences between them.

Achieving these two goals, however, will pose difficult methodological challenges. Because we are looking for typical patterns, analyses cannot rely on intuitions or anecdotal evidence. Furthermore, we need to analyze a large amount of tourist publications collected from many sources, to make sure that we are not basing conclusions on a few samples. However, with a large amount of materials, it is time-consuming to carry out the analyses and difficult to keep track of multiple language features. (Biber, 2000)

Thanks to the development of computer technologies and the advances of processing programs, “the corpus-based approach provides a means of handling large amounts of language and keeping track of many contextual factors at the same time. It therefore has opened the way to a multitude of new investigations of language use.”(ibid)

4.1.1 Corpus Linguistics Abroad

Modern corpus linguistics is often assumed to have begun in the early 1960s, but corpus-based linguistic analysis of various kinds has a long history. From the 18th century, lists and concordances of words used in the Bible were made in an attempt to show that the various parts of the Bible were factually consistent with each other. Oxford English Dictionary (OED) and Noah Webster's *An American Dictionary of the English Language* are both based on a great corpus of citations. Without electronic database and computer aids, working manually with texts of citations is hugely time-consuming and labour intensive. Many editors of OED had died before the work was completed. In the first half of the 20th century, one of the most influential corpus-based research was a pedagogical corpus compiled by Thorndike. It is a corpus of 4.5 million words from 41 different sources to make a word frequency list which was initially intended to lead to better curricula materials for teaching literacy to native speakers of English in the United States. The most important pre-electronic corpus assembled particularly for grammatical description was *the Survey of English Usage (SEU) Corpus* (Quirk, 1968). It marked the transition between earlier non-computerized corpus-based description and development of modern corpus linguistics.

Modern corpus linguistics dates far back to the era of post-Bloomfieldian structural linguistics in the U.S.A. This was the time when linguists (such as Harris and Hill in the 1950s) were under the influence of a positivist and behaviorist view of the science, and regarded the "corpus" as the primary tool of linguistics. For such linguists, the corpus—a sufficiently large body of naturally occurring data of the language to be investigated was both necessary and sufficient for the task in hand, and intuitive evidence was a poor second, sometimes rejected altogether. But there was virtually no continuity between the corpus linguists of that era. The discontinuity can be located fairly precisely in the later 1950s. Chomsky, together with his theory of TG grammar, had effectively put to fight against corpus linguistics of the earlier generation. His view on the inadequacy of corpora, and the adequacy of intuition, became the orthodoxy of a succeeding generation of theoretical linguists.

When the first computer corpus, *the Brown Corpus*, was being created in the early 1960s, generative grammar dominated linguistics, and there was little tolerance for approaches of linguistic study which are not acceptable to generative grammarians. Therefore, in 1960s, the efforts to create a machine-readable corpus of English by W. Nelson Francis and Henry Kucera (creators of *the Brown Corpus*, and now regarded as pioneers and visionaries in the corpus linguistic community) were not warmly accepted by many members of the linguistic community. The creation of *Brown Corpus* was characterized by a leading generative grammarian of the time as “a useless and foolhardy enterprise”. (Francis, 1992) However, as the first machine-readable corpus for linguistic research, *the Brown Corpus* has laid the foundation for modern corpus linguistics. More than this, its significance in corpus linguistics was further strengthened by the fact that it was compiled in the face of massive indifference from those who favored the mainstream paradigm at that time in US linguistics represented by Noam Chomsky.

4.1.2 Corpus Linguistics in China

Since the early 1970s, some Chinese researchers began to turn their attention to the special-purposed corpus compilation. At the Hong Kong University of Science and Technology, a one-million-word corpus of the English of computer science has been developed, intended to assist the teaching of English for computer science students in Hong Kong (James, 1996). The corpus consists of three 2,000-word samples from each of some 166 English language textbooks used in computer science courses in the early 1990s. Other specialized corpora with a similar applied linguistic purpose include the *Jiao Tong University Corpus for English in Science and Technology (JDEST)* and the *Guangzhou Petroleum English Corpus (GPEC)* produced in China. Both are designed to facilitate lexical analysis of particular registers, including counts of high frequency words. The JDEST Corpus was compiled in the 1980s and consists of about one million words from written English texts mainly in the physical sciences, engineering and technology. The GPEC consists of about 411,000 words comprising 700 texts from the petroleum industry from written American and British English sources of the mid-1980s. In the late 1990s, there are probably several hundred corpora compilation projects underway

internationally for various languages.

Ever since 1961, almost for four decades, corpus linguistics has gradually extended its scope and influence. "Corpus linguistics has now become a mainstream. No longer the preserve of the computer "boffin", corpus-based research is increasingly influential in many areas of language study."(Thomas, J. & Short, M., 2001). It has not revived the American structural linguists' claim of the all-sufficient corpus, but the value of the corpus, as a source of systematically retrievable data, and as a test bed for linguistic hypotheses, has become widely recognized and exploited. More important, perhaps, it has been discovered that modern corpus offers a new methodology for building robust natural language processing systems. More detailed quantitative analyses (requiring large corpora and the aid of computer) can be expected to produce results beyond the insights of native speakers.

4.2 Corpus

The notion of a corpus derived from the Latin word which means "body" can be generally considered as a large collection of spoken or written text. A corpus represents a sample of particular variety of languages in machine-readable form in modern linguistics. The following are two typical definitions of corpus:

Corpora necessarily consist of structured collections of text specially compiled for linguistic analysis, that they are large or that they attempt to be representative of a language as a whole.
(Kennedy, 2000:3)

In the language science a corpus is a body of written text or transcribed speech which can serve as a basis for linguistic analysis and description. (Kennedy, 2000:1)

Corpus is similar to an archive in form but rather different in nature. Leech (1991:11) suggests "the difference between an archive and a corpus must be that the latter is designed or required for a particular 'representative' function." The archive is only a

collection of all kinds of texts stored due to the interest of each individual. A corpus can be gathered on the basis of particular principles and special purposes, while an archive is a gathering of texts arranged randomly.

Electronic corpora can consist of whole texts or collections of whole texts. They can consist of continuous text samples or collections of whole texts; they can even be made up of collections of citations. (Kennedy, 2000)

Corpora have been compiled for many different purposes, which in turn influence the design, size and nature of the individual corpus. Generally, they are designed for linguistic research on prosody, lexis, grammar, discourse patterns or pragmatics. (ibid)

4.3 Major electronic corpora for linguistic research

We will now take a historical look and record some of the milestones of corpus linguistics.

4.3.1 The Brown Corpus

Corpus-based research is often assumed to have begun in 1961 with the advent of *Brown Corpus* (*Brown University Standard Corpus of Present-Day American English*), the first computer corpus compiled for linguistic research. "The structure of the *Brown Corpus* consists of 500 samples each of about 2,000 words of continuous written English. The resulting corpus contains approximately 1,014,300 words."(Kennedy, 2000: 24) The careful planning of the structure with the selection of text categories to represent a broad range of stylistic aspects of written American English, set a standard for corpus-based research. Another important standard was the principle of free access. (ibid)

4.3.2 The Lancaster-Oslo/Bergen(LOB) Corpus

Between 1970 and 1978 a corpus of written British English was compiled at the University of Lancaster and the University of Oslo in collaboration with the Norwegian Computing Centre for the Humanities at Bergen. *The Lancaster-Oslo/Bergen (LOB) Corpus* was intended to be a British English counterpart to *the Brown Corpus*. Like the American corpus, *the LOB Corpus* contains 500 texts of about 2,000 words each, all published in 1961, to give about one million words of running text in total. (Kennedy,

2000: 28)

4.3.3 The London-Lund Corpus (LLC)

It combines two sub-corpora together: the Survey of English Usage (SEU) Corpus and the Survey of Spoken English (SSE).

The SEU Corpus, founded by Randolph Quirk in 1959, was the most important pre-electronic corpus. The samples were half spoken and half written, of which the spoken ones “were taken from a wide range of genres including conversations, interviews, lectures, seminars, committee meetings and telephone discussions”. (Kennedy, 2000:17) The corpus was of great value except that it could not be analyzed by computer.

In 1975, however, *the SSE Corpus* was set up at Lund University in Sweden by Jan Svartvik initially to make available in electronic, machine-readable form the spoken part of *the SEU Corpus*. The original 87 texts of transcribed speech totaling some 435,000 words were supplemented by 13 more texts to produce the complete *London-Lund Corpus* of one hundred 5,000-word texts. (Kennedy, 2000:31)

4.3.4 The British National Corpus (BNC)

It is always the case that the ever-improving technology encourages more ambitious enterprises. “Between 1991 and 1995 *the British National Corpus* was undoubtedly the most ambitious corpus compilation project yet attempted.”(Kennedy, 2000:50) With 100-million-word samples of contemporary spoken and written British English, “the corpus was designed to be representative of British English as a whole and not just one particular genre, subject field or register”. (ibid) The 4,124 texts from 90% written and 10% spoken sources are widely accessible for educational, academic and commercial purposes.

4.4 Corpus Linguistics: A Brand-New Research Method

“Over the last three decades the compilation and analysis of corpora stored in computerized databases has led to a new scholarly enterprise known as corpus linguistics.” (Kennedy, 2000:1)

As a matter of fact, the term “corpus linguistics” has two meanings: First, it refers to

the research on some aspects of language with the use of corpus, that is to say, “corpus linguistics” is not ^a name of new discipline but a fresh research method. Second, it is the animadversion on the contemporary linguistic theories according to the language facts reflected by corpus, then putting forward new ideas or theories. Only on the latter level can “corpus linguistics” be recognized as a new discipline. Currently, however, this area remains untouched. Therefore, strictly speaking, corpus linguistics cannot be compared with sociolinguistics, psychological linguistics and pragmatics, etc.

Corpus linguists like all linguists are concerned primarily with the description and explanation of the nature, structure and use of language and languages and with particular matters, for example, language acquisition. Sometimes they tend to focus on lexis and lexical grammar rather than pure syntax. (Kennedy, 2000:8)

The scope of corpus linguistics involves several quite different activities. The first category is corpus designing or compiling. A second is associated with developing analysis tools of corpora. A third concerns employing corpora to describe language phenomenon. Distinguished from conventional descriptive linguistics, this corpus-based approach highlights how often particular forms are used rather than what is said or written, where, when and by whom. A fourth area of activity, regarded as the most innovative and far-reaching outcome of the corpus revolution, is applications of corpus-based linguistic description including language learning and teaching, speech recognition and translation.

4.5 The Corpus-based Approach

4.5.1 The Characteristics

What makes the corpus-based approach different? Below are the design features of this new analytical method:

“-it is empirical, analyzing the actual patterns of use in natural texts;

-it utilizes a large and principled collection of natural texts, known “corpus”, as the basis for analysis;

-it makes extensive use of computers for analysis, using both automatic and interactive techniques;

-it depends on both quantitative and qualitative analytical techniques.” (Biber, 2000:4)

4.5.2 Four Types of Corpus-based Analysis

Whether the texts have spoken or written origins, corpus-based studies have characteristically focused on four main types of analysis and description:

1 Word-based studies which explore the ecology of lexis both in terms of the occurrence and frequency of occurrence of items.the expansion of lexical studies to include collocations is likely to be one of the most innovative and productive areas of corpus-based research.

2 Studies of the co-occurrence of grammatical word-class tags as expressions of syntactic patterning and as the basis for quantitative studies of the use of syntax.

3 Studies of the co-occurrence of groups of linguistic items or processes to show by means of factor analysis the linguistic characteristics of genre.

4 Studies of the structure of discourse, especially of spoken interaction, and of the basis of cohesion in spoken and written texts.
(Kennedy, 2000:276)

4.5.3 Analysis Tools

In general, two kinds of tools are used for the corpus-based analyses: commercially available packages and specially developed computer programs. Publicly available software packages are referred to “concordancing” programs. These programs allow the user to search for specific target words in a corpus, providing exhaustive lists for the

occurrences of the word in context. They thus enable the analysis of lexical collocations (i.e., lexical-lexical association patterns), and also provide frequency information. Concordancing programs have been available for many years, and new ones join the market each year. Some of these, such as TACT and Lexa, are available at a very small or no cost.

However, concordancing programs are not made for all sorts of corpus analysis and investigations. Sometimes computer programming skills are required.

4.5.4 Statistics in Corpus Analysis

Software for corpus analysis typically provides basic descriptive statistics on the number of word forms or tokens in the corpus or section of the corpus, the number of different word types and the type-token ratios. Other descriptive statistics are sometimes provided on mean sentence length, and the number of sentences containing particular numbers of words, the number of sentences in the text, and the proportion of the sentences which have more than a specified number of words.

In this thesis, SPSS (Statistical Package for Social Sciences) 8.0 is employed as the statistical tool. Well developed and reader-friendly, it would enable researchers to test quickly and easily whether differences between observed frequencies of occurrences of items in different genres or languages are statistically significant, or whether the co-occurrence of two or more items in a sentence occurs at a greater-than-chance frequency. In conclusion, it would improve the rigour of analyses and provide a more focused direction to corpus-based analyses of language. See the application of SPSS 8.0 in Section 5.5

4.5.5 Applications

The corpus-based approach takes advantage of: computer's capacity for fast, accurate, and complex analyses; the extensive information about language use found in large collections of natural texts from multiple registers; and the rich descriptions that result from integrating quantitative findings and functional interpretations. (Biber, 2000:233) For these reasons, the corpus-based approach can be used to study a wide

variety of topics within linguistics such as individual words, grammatical features, ... author style, register patterns. In fact, it can be applied to empirical investigations in almost any area of linguistics.

Lexicography and grammar, the core areas of linguistic structure, can be studied from a use perspective by applying corpus-based techniques. For example, in the past, dictionary makers generally limited their task to identifying the possible meanings of a word. Now they can also include information about the most common uses, the frequency of related words, and the contexts in which words and meanings are most commonly found. (Kennedy, 2000)

Sociolinguistics also benefits from the corpus-based studies. In the past, the investigation of language acquisition was restricted by a small number of subjects. Now, as corpora of learners' language are compiled, studies can be based on a large number of learners, and general patterns across learners can be examined

Stylistics, too, is facilitated by the corpus-based approach. Individual authors or styles across historical periods can be investigated in a more comprehensive way than in the past, examining more texts and more language features.

Educational linguistics is no exception. The results of large-scale studies of use are helpful in designing effective materials and activities for classroom and workplace training, allowing us to help students with the language that is actually used in different target settings. (ibid)

As more and more linguistics students become familiar with using computers to address research questions, and others become comfortable with the use of more user-friendly software, we can expect the use of corpora to become more of a mainstream activity.

Chapter 5 Corpus Analysis of Tourist Texts

This chapter is the crucial part of the thesis, in which a comparative study will be conducted between the translated version of Chinese tourist publications and the corresponding English texts, based on a large database of real language examples collected on the internet. The purpose of the research is to find out the lexical and sentential features of the two varieties, in hope of providing insight into the translation of Chinese tourist publications.

5.1 The Corpus-based Approach Compared to the Conventional Approaches in Tourist Publications Translation

For more than two thousand years, translation theory was mainly limited to two traditional modes---one concerned with the study of ^{language} literature and the other with ^{literature} language. These two modes even remain dominant in today's research of translation. The first and most prominent one can be called scientific approach, "focusing on linguistic matters, rejecting any analogical solutions and subjective speculation" (Gentzler, 1993:74). This kind of translation theories is focused on the linguistic implications for two reasons: on one hand, linguistics has achieved great success in the 20th century; on the other hand, translation is defined as the transference between two languages. Consequently, translators tend to compare the two languages involved so as to find the discrepancy between the source language and the target language. Then it is justified for them to find a way to settle the difference under the guide of certain linguistic principles.

The second is the literary approach, "focusing primarily on literary concerns, rejecting theoretical presuppositions, nominative rules and linguistic jargons"(ibid). The literary approach can find eloquent expression in the translation traditions of China. Over centuries, translation has been restricted primarily to literary approach rooted in literature and philosophy. In this case, the translator's artistic sensitivity and sympathy weigh more than anything else, and consequently, literary translation tends to be willful and mutable, for every translator comprehends the text in his/her own way.

"This situation has changed during the course of the last thirty years. The study of

language has undergone radical changes: the focus of interest has widened from the purely historical to the contemporary, from the prescriptive to the descriptive, from the theoretical system to the concrete realization, from the micro-level of the sign to the macro-structure of text” (Snell-Hornby, 2001:7). The combined impact of these developments facilitated the booming of translation theories in different light. In addition, the interdisciplinary exploration of translation attracts more and more scholars, who have attempted to correlate translation with pragmatics, sociolinguistics, culture, even anthropology and so on.

Corpus linguistics is a marginal science, which arises from linguistics, computer science, cognitive linguistics and applied linguistics. It has provided linguistics with a totally new methodology through large-scale language analyses based on authentic language data. “Linguists have always needed sources of evidence for theories about nature, elements, structure and functions of language, and as a basis for stating what is possible in a language. At various times, such evidence has come from intuition or introspection, from experimentation or elicitation, and from descriptions based on observations of occurrence in spoken or written texts.”(Kennedy, 2000:7) In the case of corpus-based research, the evidence is derived directly from texts. Hereby corpus linguistics differs from the approaches to language which depend on introspection for evidence.

5.2 Theoretical Feasibility of Corpus Analysis in this Research

Any scientific enterprise must be empirical in the sense that it has to be supported or falsified on evidence and, in the final analysis, statements made about language have to stand up to the evidence of language use.

Corpus linguistics is inextricably linked to the computer, which has introduced incredible speed, total accountability, accurate replicability, statistical reliability and the ability to handle huge amounts of data. With modern software, computer-based corpora are easily accessible, greatly reducing the drudgery and boredom of dealing with the heaps of tourist materials used for the comparative analysis later. In addition to greatly increased reliability in such basic tasks as searching, counting and sorting linguistic items,

for instance, the frequency of adjectives, the length of sentences, computers can show accurately the probability of occurrence of linguistic items in text. Computers have aided the author to work with a large variety of texts and thus to seek generalizations about the distinctive features of the Chinese tourist publications and the corresponding English texts. (Kennedy, 2000)

5.3 Corpus Design in the Research

In order to conduct a cross-language comparative study, two comparable mini-corpora have been compiled, made up of written tourist texts on the Internet produced by native English writers and Chinese translators.

5.3.1 The Criteria of Sampling

How did the author pick out samples out of the sea of tourist materials? Below is the criteria adopted:

First, in terms of source, samples in this research are taken from the Internet. With the outspread of personal computers and networks, surfing on the Internet has become an indispensable part of modern life. When people plan to travel, especially foreigners, searching information on the potential destination through the Internet would be their first choice. Compared to the traditional printed products, online guides enjoy the priority of greater accessibility. Only with a click of the finger, the attractions, facilities, transportations and services of the destination will be shown on the screens, quickly and exhaustively.

Second, due to the limits of time and length of this thesis, the author focuses natural attractions texts only. 50 Chinese samples and 50 English samples (mainly from the U.S and Britain) of this category make up two corpora respectively, covering those of mountains, parks, rivers, lakes, botanical gardens, zoos, beaches and islands.

Third, the samples selected are common-tourists-oriented, excluding the formal expositions aiming at tourism experts or professionals.

Last, the size of each sample ranges from 100 to 300 words.

Based on such criteria, samples' representativeness is guaranteed and

thus enhances the reliability of the final results.

5.3.2 The Validity of Samples

The samples in the thesis are collected mainly from the Chinese official tourist websites: <http://www.cnta.gov.cn>, <http://www.qdta.cn>, <http://www.jinan.ta.com>, <http://www.visitshezheng.com>, etc. See more in Appendix A.

These samples cover such major tourist cities as Beijing, Shanghai, Shenzhen, Suzhou, Hangzhou, Qingdao, Taian, Changsha, etc. and such scenic spots as the Great Wall, the West Lake, Mount. Tai, Mount Huangshan.

The samples of English tourism publications are downloaded from the western official websites, for example, <http://www.ego.com> <http://www.cityinformation.com>, <http://www.atc.net.au>, <http://www.usa.worldweb.com>. For details see Appendix A.

The cities and scenic spots involved are New York, Washington, London, Sydney, San Francisco, Seattle, Philadelphia, the California Disneyland Adventure Park, Cave and Basin National Historic Site of Canada, etc.

5.3.3 The Proportions of Samples

50 samples of each corpus are further grouped into 10 of mountains, 10 of parks, 5 of rivers, lakes, botanical gardens, zoos, beaches and islands respectively.

5.4 Corpus Software Programs

After establishing two comparable corpora, the data will be analyzed by a computer program. According to Kennedy, analysis of a corpus by means of standard corpus linguistic research software can and frequently does reveal facts about a language which we might never previously have thought of seeking. (Kennedy, 2000)

5.4.1 The Definition of a Computer Program

Fundamentally, a computer program is a set of instructions to a computer. For our purposes, a program tells the computer how to analyze a corpus. "It tells what texts to use as input, what linguistic features to analyze and how to identify them, and what kind of output to produce. In essence, the program controls the computer." (Biber, 2000:254)

5.4.2 The Commercially Available Programs

As discussed in Chapter 4, two kinds of tools are used for the corpus-based analyses: commercially available packages and specially developed computer programs. Commercial programs mainly include: SARA of BNC, My Finder of CLEC, and Wordsmith Version 4.0. These programs allow the user to search for specific target words in a corpus, providing exhaustive lists of the occurrences of the search word in context. Frequency survey, collocation analysis, nominalization check, etc. thus can be conducted instantly with them. The latest version of Wordsmith, Version 4.0, can fulfill a wide range of tasks and can be applied to almost all corpora. It is very convenient to access Wordsmith through Internet, as is the general case to most of the other software programs developed for corpus concordancing.

5.4.3 The Awk Program

Instead of the above-mentioned programs, this research employs the awk program. Awk, named after its developers Aho, Weinberger, and Kernighan, is a programming language which permits easy manipulation of structured data and the generation of formatted reports.

The awk utility is a pattern scanning and processing language. It searches one or more files to see if they contain lines that match specified patterns and then perform associated actions, such as writing the line to the standard output or incrementing a counter each time it finds a match.

Some of the features of awk are:

- Its ability to view a text file as made up of records and fields in a textual database.

- Its use of variables to manipulate the database.

- Its use of arithmetic and string operators.

-Its use of common programming constructs such as loops and conditionals.

-Its ability to generate formatted reports. (*Awk Tutorial* from <http://www.vectorsite.net>)

The capabilities of awk extend the idea of text editing into computation, making it possible to perform a variety of data processing tasks, including analysis, extraction, and reporting of data. These are, indeed, the most common uses of awk.

Awk has two faces: it is a utility for performing simple text-processing tasks, and it is a programming language for performing complex text-processing tasks (ibid).

The two faces are really the same, however. Awk uses the same mechanisms for handling any text-processing task, but these mechanisms are flexible enough to allow useful awk programs to be entered on the command line, or to implement complicated programs containing dozens of lines of awk statements (ibid).

Awk statements comprise a programming language. In fact, awk is useful for simple, quick-and-dirty computational programming. Anybody who can write a BASIC program can use awk, although awk's syntax is different from that of BASIC. Anybody who can write a C program can use awk with little difficulty, and those who would like to learn C may find awk a useful stepping stone, with the caution that awk and C have significant differences beyond their many similarities (ibid).

5.5 The Corpus-based Operations

In this section, some of the main parameters and conventions of texts of the Chinese corpus and the English corpus are identified and compared, mainly including: type-token ratio and mean sentence length, so as to detect the differences and to provide referential norms of tourism English for the translation. Before the analysis, the author makes the following hypotheses:

- a. The type-token ratio of Chinese corpus is lower than that of English corpus
- b. The average sentence length of Chinese corpus is longer than that of English corpus

5.5.1 Type-token Ratio

a. **Word frequency.** “Count_words” is an awk program file. It tells awk to count the “total words” (tokens) and “total different words” (types) of input files, as well as the frequency of each word in the files. Here the input files refer to 50 Chinese sample texts and 50 English sample texts respectively. The following command will do this:

```
awk -f count_words.txt sample1.txt sample2.txt...sample n.txt >result.txt
```

This yields: (The following figures indicate the frequency of the words listed)

In Chinese texts:

```
...  
a 130  
in 128  
...  
China 23  
of 233  
the 344  
and 224
```

...

In English texts:

```
...  
a 198  
in 153  
...  
is 144  
of 316  
and 316
```

...

b. In the same way, the author calculated each corpus's **type-token ratio** according to different categories of natural attractions: beaches, botanical gardens, islands, lakes, mountains, parks, rivers and zoos. Input the data into *SPSS Data Editor* then we get the following tables and charts:

Case Processing Summary

	Cases					
	Valid		Missing		Total	
	N	Percent	N	Percent	N	Percent
CATEGORY * CORPUS	821	100.0%	0	.0%	821	100.0%
CATEGORY * RATIOS	821	100.0%	0	.0%	821	100.0%

Figure 5.1 The Table of Case Processing Summary

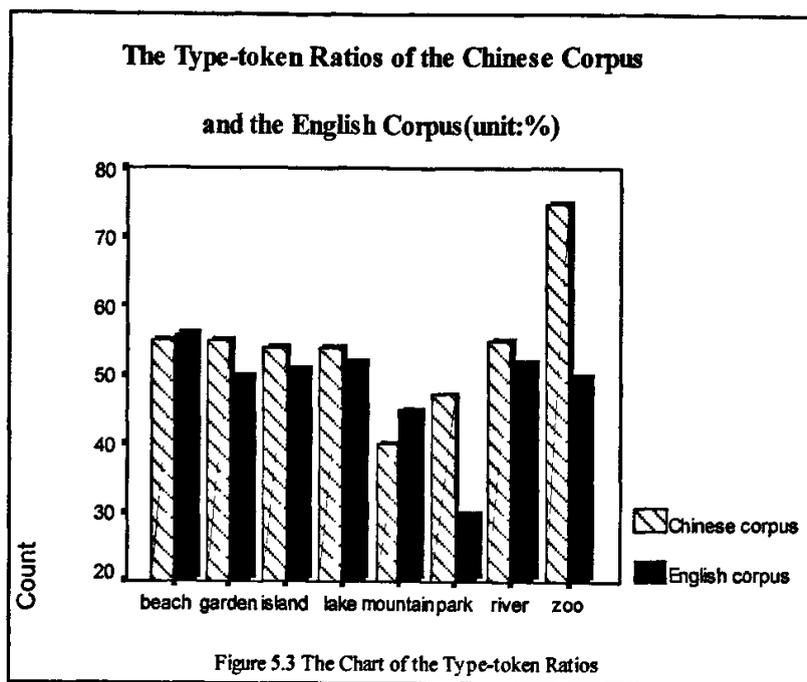
Categories * Corpora Crosstabulation (unit: %)

Count

CATEGORIES	CORPORA	
	C	E
beaches	55	56
botanical gardens	55	50
islands	54	51
lakes	54	52
mountains	40	45
parks	47	30
rivers	55	52
zoos	75	50

Footnote: C= Chinese corpus; E= English corpus

Figure 5.2 Categories * Corpora Crosstabulation



c. Findings: Figure 5.1 shows the level of validity of the data. Figure 5.2 indicates the type-token ratio of each category of natural attractions and in Figure 5.3 the data are further illustrated as clustered bars. From the results above, both similarity and differences between Chinese texts and English ones are exposed. For instance, the most “popular” words in Chinese tourist texts enjoy equally high frequency in English texts, such as “a”, “in”, “of”, “and”, most of which belong to functional words.

What is more significant, however, lies in the differences in “type-token ratios”. The bar chart shows that the Chinese corpus outdoes the English in lexical variety, especially in the category of zoo introductions, in which the type-token ratio of Chinese corpus reaches as high as 75%, compared to 50% of the English. The result is completely opposite to the author’s hypothesis. One reason, the author assumes, is that the Chinese corpus has relatively more proper nouns than the English. The introduction to *Luofu Mountain* serves as the best example.

Sample 5.1

Scenic Luofu Mountain

After an hour's ride north of Huizhou, one arrives at the Chongxu Taoist Temple on the Luofu Mountain. The Luofu Mountain is one of four famous mountains in Guangdong Province, and it stretches across the boundary between Boluo, Longmen, and Zengcheng Counties. There are wonderful things to see in these mountains, including the Chongxu Taoist Temple complex, considered to be one of the most important Taoist cave sites in China.

In the Memorial Temple to Ge Hong, there are statues of Ge Hong (284-363), sitting cross-legged and wearing a golden Taoist robe, and his wife Bao Gu. Ge Hong, in addition to being a famous Taoist sage, was also a pharmacologist and lived during the Eastern Jin Dynasty. He came to the Luofu Mountain to practice alchemy, and built the four temples Pidu, Guqing, Baihe, and Sulao; thus establishing important Taoist sacred places in Guangdong and Guangxi. The Pidu Temple had undergone repairs during successive dynasties, and in 1087, it was bestowed the name of Chongxu by Emperor Zhezong (1086-1100) of the Northern Song Dynasty.

In such a 178-word passage, the number of proper nouns reaches as many as 17, yet most of which turn out to be redundant, such as the *Pinyin* "Boluo, Longmen, and Zengcheng", meaning nothing but frustration to foreigners.

d. A further study In order to testify the author's assumption, a contrastive research is conducted between a Chinese text and an English one. Both belong to the descriptions of parks, standing out for few proper nouns:

Sample 5.2

Window of the World, Shenzhen

Window of the World is situated in the Overseas Chinese Town near Shenzhen Bay and covers a total area of 480,000 square kilometers. To enhance world culture, it integrates world wonders, historical relics, places of interest, and folk dance performances into the garden. The slogan of Window of the world is "give me a day and we will show you a splendid world". Scenic spots in Window of the World are divided in accordance with the Five Continents, forming a magnificent world together with World Square, World Sculptures Garden, International Street, and Jurassic Park. 118 scenic spots are imitated according to different scales. Visitors acclaim them as the acme of perfection because the scenes are so exquisite and vivid. Scenic spots in Window of the World are

like solid symphonic poems, while the fabulous folk performances are like moving custom paintings. When night falls, grand folk song and dance shows and carnivals parade on World Square incisively manifest the theme of window of the world-“The world will be happy together with you”.

Tokens: 176

Types: 105

Type-token ratio: 60%

Sample 5.3

Welcome to Hontoon Island State Park

This island, located in the St. Johns River in Volusia County, welcomes visitors to enjoy nature and history in quiet solitude. The island is accessible only by private boat or park ferry. Evidence of Native American habitation over thousands of years can be witnessed as visitors hike through the park. Stop in and walk through the impressive visitor center to learn more about the many inhabitants and uses of Hontoon Island over the years. Boating, canoeing, and fishing are popular activities and canoe rentals are available. Picnic areas include tables, grills, and a playground. For overnight stays, the park has rustic cabins with an enclosed sleeping space with electricity, heat, air conditioning, bathrooms, or kitchens. Overnight boat slip rentals and a tent-only campground are also available. The park's ferry operates daily from 8:00 a.m. to one hour before sunset. Located six miles west of Deland off State Road 44. For Information about Hontoon Island State Park, please call 386-736-5309.

Total words: 156

Total of different words: 114

Type-token ratio: 73%

From the samples above, it is found that the type-token ratio of the Chinese sample is lower than that of the English sample, which is contradictory to the former findings. Therefore, the author's assumption proves rational that the Chinese corpus has more proper nouns than the English corpus. They occupy part of “types” so that the type-token ratio of the Chinese corpus is higher than the English on the whole.

5.5.2 Sentence Length

a. Simply but effectively, Microsoft Word is employed to calculate the average sentence lengths of the two corpora. Since the period “.”, the exclamation mark “!”, the question mark “?”, and the ellipsis “...” signify the end of a complete sentence in English, the number of those punctuations equals the number of sentences. Based on such a presupposition, the study proceeds as the following steps:

1. Make sure all the texts are saved as Microsoft Word texts.
2. Click on Edit. Highlight Find, which can be found at the bottom of the drop-down menu. The *Find and Replace* dialog box appears.
 - (1) Select the *Replace* tab
 - (2) In the *Find what* text box, type the punctuations “.”
 - (3) In the *Replace what* text box, type the period “.”
 - (4) Click *Replace ALL* and keep record of the number of the replaced items
 - (5) Click *CLOSE*
 - (6) Deal with the punctuations “!”, “?”, “...” as the same.

Thus, the number of sentences in each text is calculated. So is the total number of sentences in the whole corpus. Last, “total number of sentences” divided by “total words” is the “average sentence length” of a corpus. Here are the results:

Corpora	Total words	Total number of sentences	The average sentence length
Chinese corpus	5507	307	17.94
English corpus	8797	568	15.49

Figure 5.4 The average sentence lengths of Chinese corpus and English corpus

Apparently, sentences in the Chinese corpus are longer than those in the English corpus on the average. Sentence length is crucial to good writing. Readers do not enjoy

long sentence after long sentence. Writers should help them “digest” ideas by varying the length of sentences.

b. A Further Study

To see the difference more directly, a case study is done between two texts of zoos.

Sample 5.4

Jinan Zoo

Jinan Zoo is a tourist attraction, where one can see many rare animals, especially China's national treasure-panda.

Jinan Zoo is located in the north suburb of the city and originally called Golden Bull Zoo. As the largest zoo in Shandong Province, it homes more than 5,000 animals covering 300 species. The layout of the zoo is based on natural gardens, which is unique in beauty. With a coverage of 1600 acres, Jinan wild Life World is located in the scenic spot “Running Horse Mountain”, south-east of Jinan and 40 km from the city. At an elevation of 700 meters, there are exuberant trees and beautiful sceneries, which make here enjoy good reputation of “Little Yunnan” in history. It was written in “A Survey of Licheng County” that “On either bank, cliff and precipice look like screen, pine and cypress are growing in the crack of the stones, flying birds are reflected in the water. Wild flowers grow in gully in spring and trees getting red in autumn.” You will enjoy yourself in the natural scenery so much with no thought of leaving.

Total of words: 185

Total of sentences: 8

The average sentence length: 23.13

Sample 5.5

Hobe sound national wildlife refuge

Refuge Facts:

- Established: 1969.
- Acres: 967 (Florida).
- Located in Martin County, FL.
- Location: the refuge headquarters is located 2 miles south of SR 708 (Bridge Road) on U.S. 1. The

refuge beach is located 1.5 miles north of Bridge Road on North Beach Road, Jupiter Island.

- Administered by Arthur R. Marshall Loxahatchee NWR.

Natural History:

- Refuge occupies some of the most productive sea turtle nesting areas on Florida east coast.
- Protects last 10 percent of unique sand pine-scrub-oak habitat remaining in Florida.
- Provides habitat for 30 threatened and endangered species including the scrub jay and gopher tortoise.
- Sand pine-scrub oak forest 232 acres, coastal dune and mangrove swamp 735 acres.
- National Natural Landmark – Reed Wilderness Seashore Sanctuary 173 acres.

Financial Impact of Refuge

- Six-person staff.
- 110,000 visitors annually.
- Current budget shared with Loxahatchee NWR (\$2,088,700 in FY 05). (157 words)

Total of words: **157**

Total of sentences: **18**

The average sentence length: **8.72**

By comparison, the Chinese sample's average sentence length is much longer than the English one. More specifically, Chinese sentences lack variety and interest. Yet, the English passage balances longer sentences with shorter ones and noun groups, and some are phrased as clear headings.

To conclude, from the statistics, we can see that type-token ratio and sentence length vary across the Chinese and the English corpus. Those differences are not obvious on the whole, but when probing into some Chinese single text, room still exists for polish.

5.6 Stylistic Differences of the Chinese Corpus and the English Corpus

As supplement of the corpus-based analysis, the author also discusses the different

styles of Chinese tourist texts and English ones. Two random samples are: the introduction of Huangshan Mountain, China and that of Grandfather Mountain, US.

Sample 5.6

Huangshan Mountain

Huangshan Mountain in east China's Anhui province is one of China's best-known scenic spots, celebrated for having all the wonders of mountain scenery characterized by spectacular rocky peaks, odd-shaped pines, crystal-clear mountain springs and seas of clouds. Known as the No. 1 Mountain Under heaven, it possesses of distinguished features. The main peak Lianhua Peak is 1,840 meters above sea level. There is plentiful rainfall in Huangshan Mountain. Therefore there are thick forests. Natural scenery in the area changes beyond prediction. Xu Xiake, a noted Chinese traveler, praised Huangshan Mountain as the best of all mountains.

In December, 1990 The Huangshan Mountain was listed in the Chronology of Recognition of World Heritages in China.

Shixin Peak in Huangshan (Seeing is Believing Peak)

Shixin Peak, in eastern Huangshan Mountain, attracts the greatest number of tourists. People say of it: The beauty of Huangshan is unbelievable until you reach this peak. That is why it is named Seeing is Believing Peak. The deep chasm below Shixin Peak is the home of the Huangshan Monkey.

Lianhua Peak in Huangshan (Lotus Flower Peak)

Lianhua Peak, the highest peak in Huangshan Mountain, is 1,860 meters above sea level. It is shaped like a lotus flower in blossom, with steep slopes and rocky peaks around. On top of the peak is a boat shaped rock named Boat for Picking Lotus Seeds. Tourists, standing here, look around as if they were reaching the sky. (239 words)

——From <http://www.cnta.gov.cn>

Sample 5.7

What is Grandfather Mountain?

Grandfather Mountain is :

- The highest peak in the Blue Ridge Mountain range (elevation 5,964 ft)
- The Mountain is privately owned and is operated as
- a scenic travel attraction

- a globally recognized nature preserve

One ticket price includes everything:

Guests purchase tickets (one ticket per person) and drive their own vehicles through the park, stopping along the way to enjoy a variety of activities.

Main Attractions

- Beautiful mountain scenery
- Mile High Swinging Bridge
- Environmental Habitats for native wildlife
 - Black Bears, Bear cubs, River Otters, Cougars, Eagles and Deer
- Nature Museum
 - excellent exhibits about the natural history of the region
 - theatre that shows nature movies made on the Mountain
 - restaurant and gift shop
- South's best alpine hiking trails
 - More than 12 miles of regularly maintained trails ranging in difficulty from easy nature walks to strenuous backcountry challenges.
- Picnicking
 - Over 100 picnic tables and grills are scattered throughout the park. The picnic area located on the right about 1/3 mile up the summit road is equipped with a water fountain and restrooms.
 - Guests are welcome to bring their own picnic. Those who have not planned ahead can purchase fried chicken or sandwiches to go from the Museum restaurant.
- Naturalist Programs
 - special activities, programs and guided tours with the Mountain's naturalist are available for an additional fee (236 words)

---From <http://www.usa.worldweb.com>

Below are the stylistic differences between the two varieties:

5.6.1 Scenery-focused vs. Programs-focused

First of all, the two texts are differentiated by focuses: Sample 5.6 lays emphasis on the description of the Huangshan's landscape, while Sample 5.7 highlights the variety of programs available in Grandfather Mountain: sightseeing, exhibitions, movies, hiking, picnicking, etc. In contrast to the colorful activities of the latter, Huangshan Mountain seems too stereotyped. The reason is that westerners and Chinese possess different

concepts, motives, objectives towards tourism. For most domestic visitors, scenery appreciation and self-relaxation always come first. For the western visitors, however, it is far to be enough. What's more, they'd like to enrich their experiences in travel. Under such a circumstance, the potential visitors are apt to have such an impression that they can do nothing but sightseeing during the Huangshan tour.

5.6.2 Flowery vs. Plain

Apparently, Sample 5.6 is more flowery or ornate than Sample 5.7. Below are the details.

a. Frequency of Adjectives

In Sample 5.6, there are altogether 16 adjectives out of the total 239 words, such as "best-known", "celebrated", "spectacular", "highest", "distinguished", "noted", etc. In addition, these adjectives, confined to the descriptions of being beautiful or famous, are repeatedly used. This kind of tautology is not effective to highlight the uniqueness of the place in question and thus will no doubt fail to arrest the attention of the potential tourists. In Sample 5.7, only 7 adjectives appear out of the total 236 words, such as "highest", "globally recognized"; "beautiful", "excellent", "special", etc. These adjectives and adverbs themselves are just descriptive ones, not as flowery as "best-known", "celebrated", "spectacular". A general rule of English writing is to write with nouns and verbs, since they are the most forceful words. But it is Chinese literary tradition to pursue the ornate style of the language in any type of writing. They try their best to avoid plain words, racking their brains to make the text beautiful.

b. Figurative Use of Language

For example, in Sample 5.6, the parallel construction "...rocky peaks, odd-shaped pines, crystal-clear mountain springs and seas of clouds", simile in "It is shaped like a lotus flower in blossom, with steep slopes and rocky peaks around", hyperbole in "Tourists, standing here, look around as if they were reaching the sky." In contrast, such rhetoric patterns are not found in Sample 5.7. Even though figurative use of language is exploited by both Chinese and English writers in many situations, English writers use it less frequently than the Chinese people. If the Chinese text culture contains many similes

and hyperboles, it may set up many vivid images in the Chinese readers' mind, but sound unnatural or even absurd to English tourists.

c. Allusion

Allusions can be easily spotted from the Chinese tourist materials. Two examples stand out from Sample 5.6: “Xu Xiake, a noted Chinese traveler, praised Huangshan Mountain as the best of all mountains.” “In December, 1990 The Huangshan Mountain was listed in the Chronology of Recognition of World Heritages in China.” The wide use of allusions reflects Chinese’s deductive thinking pattern. And they are used to thinking on the basis of a preset conception. They revere authority, whose words will largely influence their behaviors and feelings.

English-speaking natives, on the opposite, are of analytical and inductive thinking pattern. They do not respect authority as much as Chinese do. Instead, they pursue the precision in detail and tend to respect “facts”.

5.6.3 Uninformative vs. Informative

In Sample 5.6, readers can catch nothing but a vague image of the mountain. Other necessary information such as accommodation, restaurant, transportation, expenditure are not mentioned at all. As a result, after reading it through, readers still have no idea of how to approach there.

Nevertheless, catering for the real needs of tourists, Sample 5.7 is merited with all-round information, including the location of the mountain, the way to access it, main attractions and entertainment facilities as well. Moreover, the information in English tourism publications is usually classified into many titled categories clear and helpful to readers as displayed in Sample 5.7:

Grandfather Mountain is:

One ticket price includes everything:

Main Attractions:

5.6.4 Formal vs. Informal

Sample 5.6 is very formal in style. Its formality lies in its formal words and dull

sentence structure. Take the first paragraph as illustration. The beginnings of the total seven sentences are as the following list:

- a) Huangshan Mountain... is...
- b) Known as the No. 1 Mountain Under heaven, it possesses...
- c) The main peak Lianhua Peak is...
- d) There is...
- e) Therefore there are...
- f) Natural scenery in the area changes...
- g) Xu Xiake... praised...

In the seven sentences, Sentence (1) and (3), Sentence (2), (6) and (7), Sentence(4) and (5) share the same structures. Consequently, the mountain appears stiff-faced to the potential tourists who are interested in it.

The informality of Sample 5.7 is couched in the appealing sentences, colloquial and casual expressions, etc., which endows readers with an easy and comfortable experience. It starts with an interrogative sentence so as to arouse the curiosity of readers. Instead of complete sentences, clear-cut phrases play the dominant role, which impress readers by conciseness and economy, for example, "a scenic travel attraction", "a globally recognized nature preserve". The informal tone is also expressed through a number of agreeable suggestions for tourists, such as "Guests purchase tickets (one ticket per person) and drive their own vehicles through the park, stopping along the way to enjoy a variety of activities." "Guests are welcome to bring their own picnic." The frank, informative and suggestive style impresses readers as if he/she were facing a friendly guide who is presenting the mountain's attractions and offering some tips.

5.6.5 Inaccessible vs. Accessible

As to the degree of accessibility, the above two samples also can serve as best illustrations. In reading the introduction to *Grandfather Mountain* in Sample 5.7, Chinese readers may meet nothing inaccessible, while the English-speaking tourists would not be so lucky in reading the introduction of *Huangshan Mountain* because of the *Pinyin* like

“Huangshan”, “Anhui”, “Xu Xiake”, “Shixin”, “Lianhua”.

To sum up, English tourist publications like Sample 5.7, instructive and credible, would be impressive and profitable for readers who want to have a clear idea about the locations of the attractions and the way to approach them. This sample represents the common features of English tourist publications, such as informal diction, informative contents, succinct style, agreeable tone, accessible information. The case analysis of the introduction to Grandfather Mountain is intended to provide some valuable norms and criteria for the translation of the Chinese tourism culture. For instance, the introduction to Grandfather Mountain is very clear and impressive to readers, for it tells all the necessary information in detail and puts forward agreeable and helpful suggestions.

Chapter 6 Proposed Translation Methods for Tourist Publications

As discussed in Chapter 2, the ultimate purpose of the Chinese tourist publications is to motivate and persuade foreigners to make their decision to buy the tourist product; that is, to visit China. Therefore, as the main channel, the translated tourist publications must be acceptable, readable and appealing to them. Yet, in Chapter 4, it is observed that there do exist some differences between Chinese texts and English materials, such as some linguistic features and style, which may pose barriers to visitors' understanding. Bad instructions are bad for business. Thus, translators' task is to bridge those gaps and make some adjustments to cater for foreigners. In this chapter, a series of strategies are proposed in response to those differences, in the hope of facilitating to produce better tourist texts.

6.1 Strategies on Diction

6.1.1 Devote Special Effort to Producing Plain English

Similes, metaphors, hyperboles and allusions are easily spotted in Chinese tourist publications, not to say the ocean of adjectives and adverbs boosting scenic sites. Instead of being so flowery, English texts enjoy a plain style.

Plain English refers to: the writing and setting out of essential information in a way that gives a co-operative, motivated person a good chance of understanding the document at first reading, and in the same sense that the writer meant it to be understood. This means pitching the language at a level of sophistication that suits the readers and using appropriate structure and layout to help them navigate through the document. It does not mean always using simple words at the expense of the most accurate words or writing whole documents in kindergarten language. (*Plain English* from <http://www.askoxford.com>)

Plain English is an essential tool for effective communication. Information is clearer and more understandable when presented in straightforward terms.

Here are some top tips for keeping the tourist texts user-friendly:

- (1) Remember the readers – Usually readers haven't visited the destination before, that's why they're reading the introductions.
- (2) Use words readers are likely to understand.
- (3) Use only as many words as you really need.
- (4) Prefer the active voice unless there's a good reason for using the passive.
- (5) Use the clearest, crispest, liveliest verb to express your thoughts.
- (6) Use vertical lists to break up complicated text. – Readers waste time and make mistakes if the information they need is buried in long paragraphs.
- (7) Put your points positively when you can.
- (8) Reduce cross- references to the minimum.
- (9) Try to avoid sexist usage.
- (10) Put accurate punctuation at the heart of your writing.
- (11) Organize the materials in a way that helps readers to grasp the important information early and to navigate through the text easily.
- (12) Consider different ways of setting out the information.
- (13) Spare no effort to producing lucid and well-organized texts
- (14) Use clear layout to present plain words in an easily accessible way.

6.1.2 Keep the Use of *Pinyin* under Control

According to the statistics in Chapter 5, the Chinese corpus's type-token ratio is higher than the English corpus in general, but comparing two single texts with few proper nouns, the result is opposite. It proves that proper nouns occupy part of types in the Chinese corpus. No doubt that proper nouns are overused in Chinese tourist texts. Take Sample 5.1 (*Scenic Luofu Mountain*) as illustration. A dozen of names of places and persons are ready to be challenged. Even a Chinese citizen will find those *Pinyin* inaccessible and confusing, let alone the foreigners. What's worse, most of the proper names have no point at all.

So, avoid employing *Pinyin* unless you have to. Try to reduce alien elements in writing, which may pose great threats to communication.

6.1.3 Maintain the Consistency of Proper Nouns

Another problem to be solved is that a certain scenic spot or historical site may have quite different versions of its English name. For instance, “太湖” is translated into “Tai Lake” and “Taihu Lake”. “黄山” has renderings of both “Huangshan Mountain” and “Huang Mountain”. The famous garden “豫园” in Shanghai, also has two English names “Yu Garden” and “Yu Yuan Garden”. Some other inconsistent names of historical sites the author has collected are listed below:

敦煌莫高窟: Mogao Grottoes in Dunhuang / Dun Huang Buddhist Cave.

周口店北京猿人遗址: The Site of “Peking Man” at Zhoukoudian Village / The Cave of “Peking Man”/“Beijing Man” Site/Zhoukoudian-Home of Peking Man.

保叔塔: Tower of Blessing King Shu/Baochu Pagoda/ Baochu Ta/King Qian Chu Pagoda.

紫禁城养心殿: Palace of Culture of the Mind/Hall of Mental Cultivation. (和珊, 2004)

The list is far from complete. Obviously, the discrepancies of proper nouns will often puzzle and frustrate foreigners. In addition, it can also make negative effects on the publicity of these tourist attractions.

6.2 Strategies on Sentence Length

As compared in the previous chapter, sentences in Chinese tourist texts are longer on the average than those in English. Sentence structure, in addition, appears dull and stereotyped.

How long can a decent sentence be? It depends on the amount of information, the coherence of its grammatical elements, and the capacity of the reader's attention and memory span. Although some lengthy sentences are easy to comprehend, sentence length places a burden on the reader's short-term memory. Confusion may result — or irritation.

Short, clear sentences are a hallmark of plain English. Almost everything written by good writers has an average sentence length between 15 and 20 words. This does not mean writing every sentence the same length. Good writers naturally vary the length and rhythm of their sentences. They balance longer sentences with shorter ones, but they keep their average sentence length below 20 words. (*Keep your essay sentence length under control* from <http://www.write-an-essay.com>)

Follow this advice:

- Keep the average sentence length between 15 to 20 words,
- cut down long sentences, and
- vary the length of sentences

An average of between 15 to 20 words for essays and research papers helps produce a clear and readable style. It's the same average as well-respected publications such as *New Scientist* or the *Economist* consistently produce. Below 15 words and there's a danger your writing will become too choppy and disjointed. As you go above the

20-word average, the text becomes more difficult to read. An average sentence length of 25 words is far too high. (ibid)

Make sure not to allow long sentences — over 40 words — to creep into the tourist text. Faced with long sentences, readers often give up halfway through or forget the start of the sentence by the time they reach the end. The more words in a sentence, the harder it becomes to understand. By comparison, it is difficult to write an unclear sentence if it is short. Edit long sentences by cutting out wasted words or by breaking them into two or more shorter sentences.

Finally, translators need to vary their sentences. A sentence length average of 18 words is fine, but if 90 percent of the sentences are 16, 17, 18 or 19 words long, the style will lack variety and interest. Translators should balance longer sentences, say up to 35 words with shorter sentences, some under five words, some phrased as direct questions, and still get an average sentence length of between 15 to 20 words.

6.3 Strategies on Stylistic Differences

We cannot talk about stylistic differences between English and Chinese texts of tourism culture, without mentioning the differences in tourism culture between the English-speaking nations and China. Tourism culture, in short, is the summation of material wealth and spiritual wealth whatever is pertinent to tourism. Every CULTURE is different from another. As a sub branch of CULTURE, tourism is undoubtedly unique, which is featured by a country's distinctive nationality, including mode of thinking, ethics and values, aesthetics, religion and customs. That's why westerners and Chinese possess different concepts, motives, objectives towards tourism. A Chinese learner or user of English, unaware of these cultural differences, may produce a description in English that sounds unnatural or absurd to native speakers.

6.3.1 Motives

For most domestic visitors, the most important reason for their touring is to admire

the beautiful scenery and relax. However, for the western visitors, it is not the case. They come to distant China not just for the natural scenery, but the mysterious oriental culture. The latter is more important compared with the former. They are curious about the culture of China, for example, the customs, life styles, history etc. So, when we advertise for the scenic spot to the westerners, we should be aware of the different tastes of the tourists. We should have different emphasis or focus in order to satisfy the need of different tourists.

6.3.2 Mode of Thinking

As discussed in Section 5.6.2, English native speakers are of analytical and inductive thinking pattern. They do not respect authority as much as Chinese do. In contrast, Chinese people are of deductive thinking pattern, and they revere authority of famous persons. The words of famous persons will largely influence their behaviors and feelings. A lot of allusions are used in tourist texts to cater for the Chinese tourists' interest. Here is another example.

Sample 6.1

Famous Men's Birthplaces

Woyang County in west Anhui was hometown of the ancient Chinese thinker Lao Zi. A local temple is dedicated to him, and Huagu Pavilion was built during Kangxi reign of the Qing. Haozhou was the birthplace of such famous Three Kingdoms men as Cao Cao and Hua Tuo. The tombs of the Caos, and the temple in honor of Hua Tuo, are still there. The tomb of Li Bai, the celebrated Tang poet, and the Taibai Pavilion dedicated to him, are found in Ma'anshan, where he died during a visit. Ouyang Xiu, a major Song, writer, wrote his immortal "Accounts of the Drunkard's Pavilion" at Langya Mountain in Chuzhou. Zhu Yuanzhang, the founding father of the Ming, was born in Fengyang County, where ruins of his tomb can still be found.

-----From [http:// www.cnta.com](http://www.cnta.com)

6.3.3 Values

Value should be a kind of common criterion that people use to measure and judge their words or deeds. There are many contrasts between eastern and western values. Among them, the contrast between individualism and collectivism is one of the most fundamental. Individualism, the core of westerners' value system, refers to a certain

political and social philosophy that emphasizes personal liberty and freedom from outside bounds, with certain egoistic meaning. Other components of individualism include self-heroism, liberalism, and in some extreme cases even anarchism. Collectivism, the typical of Chinese, on the other hand, regards an individual as a mere member of a group. One's personal interests should be put in a secondary place since the realization of collective interests is an essential prerequisite of almost every thing. When personal interests are in conflict with collective ones, it is usually the former that has to give away. Consequently, discipline, cooperation and unity are especially emphasized. In some scenic spots, there are many items for physical training, challenge the visitor's physical limits or courage. When we introduce these activities to the westerners, we'd better add something to emphasize the value of the activities—challenge oneself, challenge the difficulties, and to be a hero! Since it is consistent with the individualism of the westerners, it arouses much interest of them. (卢桂华, 2002)

6.3.4 Aesthetic Values

The factors of natural environment, social context, politics, economy and religion will greatly influence the aesthetic values of a person. What is arresting to Chinese may not be so appealing to westerners. For example, a scenic spot is advertised as “十里蛙声不断,九溪曲流屏屏” (ibid). For the Chinese, such a description is quite desirable, quiet and beautiful. However, for the westerners, they have different response. In their eyes, it is a quite noisy place, let alone they like it. Only when we know the differences between the different readers, can we produce the English version more acceptable and attractive.

6.3.5 Religions and Customs

Religion has placed lasting influences through human history. Both western and Chinese people have a variety of religious beliefs that are part of their respective cultures. Thus many scenic spots and historical sites are of a strong religious flavor. For westerners, all these religions are full of oriental mysteries and enchantment since their own are mainly based on western dominant religion—Christianity. A large number of tourist materials involve this culture for foreigners to experience a totally different religious

atmosphere. As a result, when the original text mentions some terms of religion and traditions, we'd better make some necessary explanation.

Sample 6.2

In the mornings or evenings full of clouds and mist, standing on the high peaks, and viewing in the same direction as the light, tourists may see a colorful halo with blue color inside and red color outside present on the dimly discernible mist veil. In the middle of the halo, a human figure is reflected, and it looks just like the colorful halo over the head of the Buddha, so it is called "the Buddhist Light" or "the Treasure Light". The Buddhist Light of Taishan Mountain is a kind of natural phenomenon of optical diffraction, which mostly appears under the conditions of a half sunny and half misty day during the period from June to August, and at the time when the sun shines slantingly.

——From <http://www.cnta.com>

In addition, the influences of customary differences are quite common. For instance, foreigners can hardly habituate themselves to such questions raised by the Chinese as "How old are you?" "What about your salary?" etc. In fact the so-called human interest of the Orientals is a vital part of Chinese traditional culture. It is necessary for the Chinese to express their concern over others and to take human relationships seriously. But to the Westerners, those questions seem to interfere with others' privacy. Nowadays, however, many foreigners change their minds when they become aware of this habit of the Chinese people.

Take pines as example. When a Chinese visit ^{Mount. Huang} Huangshan (黄山), seeing the different and graceful postures of pine, it is quite natural for him to be reminded of the noble qualities of pine: persevere, firm, resolute etc; and then the unflinching will of a person. But, for a westerner looking at the pine, that association hardly occurs. Just like the church means a lot to the westerners, but nothing to most Chinese. If we want the western visitors to have better understanding of the Chinese culture, we can add some cultural meaning of the word "pine" in Chinese to the introduction.

These aspects are just three most typical ones of tourism culture, while many more are involved in tourist publications like history, cuisine, language, etc. All these aspects reflect Chinese particular culture and become a major stimulus to attract foreigners. Meanwhile, full of special cultural terms, they also bring great challenges to translators.

Conclusion

This thesis has delved into tourist publications translation from a perspective of corpus linguistics. This corpus-focused approach goes beyond the constraints of traditional linguistic translation study and has opened a new window to the study of translation.

The thesis has been intended to compare the lexical, sentential and stylistic features of Chinese tourist texts and English source texts, and analyze the differences between them. The ultimate purpose is to propose effective and feasible strategies to bridge the gap and produce more qualified texts. By comparison to other researchers, the author adopts corpus linguistics as the main theoretical framework. As a mainstream now, corpus linguistics has provided linguistics with a totally new methodology through large-scale language analyses based on authentic language data. Linguists have always needed sources of evidence for theories. Previously, such evidence has come from intuition, introspection, experimentation or elicitation. Yet, in corpus-based approach, the evidence is derived directly from texts, which paves the way to more reliable study on the differences in certain linguistic items between two languages or language varieties. That's why the author applies this new method to tourist translation.

With reference to corpus linguistics, two mini-corpora of Chinese tourist texts and English ones have been constructed, totaling 100 samples. Then type-token ratios and mean sentence length of each corpus are examined by awk program and SPSS software. The result turned out that the type-token ratio of Chinese texts is higher than that of English, partly because of the overuse of proper nouns in the former. As for the average sentence length, Chinese is a bit longer than English. In sentence structure, however, English is much more flexible and variable than Chinese. Those analysis and illustrations on relatively large database are sure to be reliable and helpful for improving the quality of the translated version and finally achieving a better publicity effect and appealing to more foreign tourists.

Apart from the corpus-based analysis, the author also devotes special effort to the

stylistic differences between the Chinese texts and the English through a case study. Due to the different cultural background, the two variety^s of tourist writing differ greatly. In acceptability and readability, Chinese publications still have a long way to go, especially in diction.

The last chapter of the thesis focuses on strategies of translation in response to those differences. Three aspects are involved: diction, sentence length and style. On diction, the author suggests a plain style, that is, using more nouns and verbs, less adjectives and adverbs. In respect to *Pinyin*, the author advises deleting those pointless proper nouns so as to enhance accessibility. On sentence length, clear-cut headings, short sentences and rhythmical patterns are advocated. As for stylistic differences, the author believes that it is deeply rooted in culture. Only with strong cultural awareness can translators produce more foreigners-friendly texts.

The study is of considerable practical value, for it broadens the horizon of the related research and provides a relatively more scientific perspective to Chinese tourist publications translation. However, due to the limited time, the author only discusses written tourist texts. In fact, tour guides' oral introductions are also an indispensable part of tourist translation. In addition, in corpus-based research much more needs to be explored, for example, frequency survey, collocation analysis. Therefore, the author calls on an earlier establishment of a large-scale corpus of tourist texts.

Bibliography

- Biber, Douglas. et al. *Corpus Linguistics* [M]. Beijing: Foreign Language Teaching and Research Press. Cambridge University Press. 2000.
- Castello, Erik. *Tourist Information Texts: a Corpus-based Study of Four Related Genres* [M] Italy: Unipress. 2002.
- Francis, W. Nelson. *Language corpora BC* [M]. In Svartvik. 1992, 17-32.
- Gentzler, Edwin.. *Contemporary Translation Theories* [M]. London and New York: Routledge. 1993.
- Goeldner, C.R. Ritchie, J.R. and McIntosh, R.W. *Tourism, Principles, Practices, Philosophies* [M]. John Wiley & Sons, Inc. New York. 2000.
- James, Gregory. *English in Computer Science* [M]. Hong Kong: Longman. 1996.
- Kennedy, Graeme. 2000. *An Introduction to Corpus Linguistics*. [M] Beijing: Foreign Language Teaching and Research Press.
- Leech, Geoffrey. "The state of the art in corpus linguistics", in Aijmer K. and Altenberg B. (eds.) *English Corpus Linguistics: Studies in Honour of Jan Svartvik*. London: Longman. 1991,8-29.
- Corpora and Theories of Linguistic Performance* [M] In Svartvik. 1992 b, 105-22
- Macdonald, Janet. *Travel Writing* [M]. UK: St Edmundsbury Press. 2002.
- Oxford Advanced Learners Dictionary of Current English*. Oxford: Oxford University Press. 1997, 1167
- Quirk, et al. *A Grammar of Contemporary English* [M]. London: Longman Group Ltd. 1968.
- Sanz, Rosa Lore's. *The Translation of Tourist Literature: the Case of Connectors* [J]. *Linguistics' Communications*. 2003, 22(3): 291-308
- Snell-Hornby, Mary. *Translation Studies: An Integrated Approach* [M]. Shanghai: Shanghai Foreign Language Education Press. 2001.
- Thomas, Jenny and Short, Mick. *Using Corpora for Language Research* [M]. Beijing: Foreign Language Teaching and Research Press. 2001.
- Awk Tutorial* <<http://www.vectorsite.net>>
- Keep your essay sentence length under control*. <<http://www.write-an-essay.com>>
- Plain English*. <<http://www.askoxford.com>>
- 安春平. *旅游文本中文化负载词的符号学翻译研究*[MA]. 大连:大连理工大学, 2005.

- 卜绍先. 民族历史文化与对外旅游翻译[J]. 大理师专学报, 2000, 48 (4):72-75.
- 陈白颖. 旅游英语及其翻译[J]. 浙江树人大学学报, 2003, 3 (4): 55-85.
- 陈代球. 汉语旅游文本中文化负载及意象的翻译:功能翻译观.[MA] 广东:广东外语外贸大学. 2005
- 陈刚. 导游翻译中的文化背景和心理因素[J]. 中国翻译, 1987(3):26-29.
- 陈刚. 涉外导游词翻译的特点及策略[J]. 浙江大学学报(人文社科版), 2002, 32 (2):67-73
- 陈红. 旅游资料翻译技巧探析[J]. 信阳师范学院学报, 2003, 23 (4):88-90.
- 陈淑华. 英汉修辞与翻译[M]. 北京:北京邮电学院出版社, 1990
- 段连城. 呼吁:请译界同仁都来关心对外翻译[A]. 中译英技巧文集. 北京:中国对外翻译公司, 1992.
- 傅文. 文化内涵与旅游翻译浅论[J]. 山东农业大学学报, 2001, 3 (3):80-81.
- 高金岭. 从中西审美方式的差异看旅游材料中景物描写的翻译[J]. 山东外语教学, 2003(3)
- 桂诗春, 宁春岩. 语言学方法论 [M] 北京:外语教学与研究出版社. 1997
- 郭瑞娟. 文本类型理论与旅游资料英译[J]. 郑州航空工业管理学院学报, 2003 22(4):98-100.
- 和珊. 旅游宣传材料中文化因素的功能翻译. [MA] 西安:西安电子科技大学. 2004
- 胡壮麟. Linguistics: A Course Book [M]. Beijing: Peking University Press. 1988.
- 纪俊超. 旅行社外语导游汉译英问题研究[J]. 海南大学学报人文社会科学版 2002. 20 (3): 96-100.
- 贾顺厚. 英译旅游材料中的文化处理[J]. 山西大学学报, 2002: 25 (4): 90-92
- 贾文波. 从汉英景物描写民族审美差异[A]. 杨自俭. 《英汉语比较与翻译》. 上海外语教育出版社, 2000: 230-41
- 靳海林. 美学与旅游作品的英译[J]. 外国语, 1993 , (2): 38-41
- 金惠康. 广东旅游翻译探讨[J]. 上海科技翻译, 2003, 20 (2): 20-24.
- 柯细香. 浅谈 Skopostheory 与旅游资料的英译[J]. 中南民族大学学报. 2002, 22(2): 90-91.
- 赖余. 浅谈旅游资料的英译[J]. 中国翻译, 1986, No. 1: 23-25
- 廖洪中. 江西旅游景点推介英译初探[J]. 南昌大学学报, 2002, 33 (3):121-126.
- 刘和林. 旅游景点山水诗的英译艺术[J]. 湖南农业大学学报, 2003 (2):62-64.
- 卢桂华. 旅游材料翻译中的文化现象[MA]. 重庆:重庆大学, 2002.
- 蒲元明. “谈谈旅游资料英译中的几个问题” [J]中国翻译, 1987, (3):30-31
- 苏冰. 从文化翻译观谈中国旅游文本的英语翻译. [MA] 山东:山东大学. 2003

- 苏丽琴. 从功能角度看旅游资料英译[J]. 上饶师范学院学报, 2001, 21(1): 85-87.
- 王丽秀. 旅游信息文本的题材分析: 对于 ESP 写作教学的启示: [MA] 广东: 广东外语外贸大学, 2002.
- 文军, 邓春, 辜涛, 蒋宇佳. 信息与可接受度的统一[J]. 中国科技翻译, 2002, 15(1): 49-64.
- 向萍主编. 旅游英语[M]. 北京: 高等教育出版社, 1998
- 邢怡. 试论旅游宣传资料翻译的标准和改进翻译质量的建议[J]. 旅游科学, 2001 (3): 39-42.
- 熊向阳. 少数民族风情译名的体会及困惑[J]. 黔东南民族师专学报, 1998, 16(2): 48-50.
- 杨端和. 语言研究应用 SPSS 软件实例大全. 中国社会科学出版社, 2004
- 杨寿康, 刘慧梅. “从文化角度看旅游资料的英译” [J] 中国翻译, 1996, (5): 10-14
- 叶苗. 旅游景点英译名初探[J]. 温州师范学院学报, 1995, 1 (4): 36-40.
- 一舟. 旅游英语翻译偶得[J]. 广西师范大学学报, 2002 (1): 101-104.
- 姚宝荣, 韩琪. 旅游资料英译浅谈[J]. 中国翻译, 1998, (5): 27-30
- 印晓红. 英语导游翻译的语言艺术与技巧[J]. 吉首大学学报, 1997, 1 (2): 64-65

Appendix A

The websites from which samples are downloaded:

<http://www.cnta.gov.cn>

<http://www.qdta.cn>

<http://www.jinan.ta.com>

<http://www.visitshenzhen.com>

<http://www.bjta.gov.cn>

<http://www.scta.gov.cn>

<http://www.xmtravel.com>

<http://fjta.com>

<http://visitgd.com>

<http://tourhn.com>

<http://www.usa.worldweb.com>

<http://visitbritain.com>

<http://www.cityinformation.com>

Appendix B

The programs of *count_words*:

```

BEGIN {FS="^[A-Za-z]" }           #以所有的非字母符号作为记录分界符。
    {
        for (i=1;i<=NF;i++)      #用 for 循环语句建立数组。
        {
            if ($i != "") count[$i]++      #为每一个词计数。
        }
    }
END {
    for ( i in count )           #for ( i in array)语句对数组中的每一个成员进行操作。
    {
        print i "\t" count[i]      #打印单词及出现频率。
        total+=count[i]           #计算总字数。
        words++                   #计算总词汇量。
    }
    print "Total words\t" total      #打印总字数。
    print "Total of different words \t" words #打印总词汇量。
}

```