

# 基于语料库的旅游文本 英译文词汇特征及翻译研究

熊 兵

(华中师范大学 外国语学院, 湖北 武汉 430079)

**摘要** 本文基于汉英旅游文本平行语料库和英文旅游文本可比语料库,对旅游文本英译文的词汇特征及特定旅游词汇的汉英翻译问题进行了探讨。研究发现,旅游文本英译文与旅游文本英文源文的词汇难度非常接近,前者比后者略低,但并无显著差异。与旅游文本英文源文相比,旅游文本英译文的词汇变化不够丰富,词汇重复率偏高,词汇密度略低,其信息量也略低一些,第一人称和第二人称代词的使用频率明显偏低。旅游文本英译文中,对自然景观特别是山川河流的介绍占重要分量;而旅游文本英文源文中,对城市景观及旅游设施的介绍占很大比重。在汉英旅游翻译中,译者需加强对某些最常见的共核动词(如 take, get)的灵活使用,并要特别注意对名词“home”及动词“offer”的灵活运用。另外,“scenic spot”并非“景点”一词理想的译文,“tourist spot”、“travel spot”及“beauty spot”比“scenic spot”更适合作为“景点”一词的对应译语。

**关键词** 语料库; 旅游文本英译文; 词汇特征; 翻译

## 一、引言

传统上,对旅游文本的语言特征及其翻译的研究,多是基于某种翻译理论(如功能对等理论、目的论等),同时也基于研究者对旅游文本的印象式或感悟式的个人体验,以此探讨旅游文本的语言特征和翻译方法与技巧,如陈刚<sup>①</sup>,方梦之、毛忠明<sup>②</sup>,贾文波<sup>③</sup>,伍锋、何庆机<sup>④</sup>,等等。近些年来,随着语料库翻译学的发展,学界开始把语料库的理论与方法引入到旅游文本的翻译研究。这一新的研究视角拓宽了旅游翻译的研究渠道,增强了研究的实证性、客观性和可验证性,促进了旅游翻译研究的发展,其中较有代表性的研究者包括李德超、王克非<sup>⑤</sup>、侯晋荣<sup>⑥</sup>、肖庚生、陈欣<sup>⑦</sup>、张苇、韩江洪、夏蓓洁<sup>⑧</sup>。不过,这些研究多泛泛讨论汉英平行语料库的构建及旅游文本的各类语言特征,对英译旅游文本的词汇特征及该特征与原创英文旅游文本词汇特征的异同挖掘不够深入,而这些特征与汉英旅游

翻译的关系也尚未得到明晰阐述。有鉴于此,本研究拟基于笔者自建的“旅游文本汉英平行语料库”及“旅游文本英文可比语料库”,对英译旅游文本的词汇特征、该特征与汉英旅游翻译的关系以及特定旅游词汇的翻译问题进行深入考察,以期能深化对汉英旅游文本词汇特征的认识,并对汉英旅游翻译提供一定的借鉴和指导。

## 二、语料、研究工具和研究思路

### (一)语料

本研究基于笔者所构建的旅游文本语料库下的两个子语料库。第一个子语料库为“旅游文本汉英平行语料库”,从该库中,我们提取了经过句对齐和词性赋码处理的汉英双语平行语料共计 191,940 字/词,其中汉语原文语料(corpus of Chinese tourism text, 简称为 CCTT)109,431 字,其英译文 82,509 词,此英译文语料在本研究中称为“旅游文本英译文语料”(corpus of translated English

收稿日期 2016-05-15

基金项目 教育部人文社会科学基金项目“应用文汉英双语平行语料库的构建及其应用研究”(10YJA740104);华中师范大学人文社科研究基地湖北文化海外传播研究中心开放式基金项目“基于语料库的湖北旅游文化对外译介研究——一个文化比较的视角”(HBWHHWC201601)

tourism texts,简称 CTETT)。这些语料内容涉及旅游景点介绍及游览(如故宫介绍、湖北旅游、长江三峡旅游、西安旅游等)、旅游指南、旅游服务(如餐饮、酒店预订、公交服务)等不同主题,英译文翻译质量较好。

第二个子语料库,为“旅游文本英文可比语料库。”这是个单语(英语)可比语料库,由“旅游文本英译文”和“旅游文本英文源文”构成。其中“旅游文本英文源文”(即原创英文旅游文本)大都采集于国外的英文旅游网站和旅游手册,具有较好的语言质量。其内容涉及景点介绍及游览(如大峡谷景点介绍、旧金山旅游介绍)、旅游指南、旅游服务(如餐饮、酒店预订)等方面。从该库中我们提取经过词性赋码处理的“旅游文本英文源文语料”共计 82,401 词(corpus of original English tourism texts,简称 COETT),该语料与前述 CTETT 一同构成可比语料库的一组研究对象。这两组语料总形符数十分接近,便于研究中(必要时)对相关数据进行直观的比较。

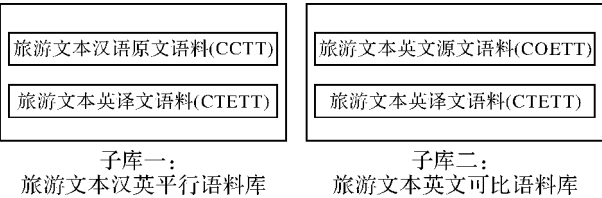


图 1 本研究的两个子语料库

(二)研究工具

研究工具主要为 WordSmith (6.0) 和 ParaConc(1.0)。

WordSmith 是一个用于观察文字在文本中表现的语料分析软件。该软件由英国利物浦大学 Mike Scott 设计,牛津大学出版社出版。自该工具

开发以来已经过多次升级,功能不断得到完善,界面更加友好,目前最新版是 WordSmith Tools 6.0。其中主要包含 Concord (语境共现检索工具)、WordList (词频列表检索工具)、KeyWords (关键词检索工具)等工具。

ParaConc 是一个用于双语平行语料库检索的软件。该软件由新西兰奥克兰大学应用语言学系教授 Michael Barlow 开发,目前版本为 1.0。该软件主要用于双语句对齐语料的检索,可以方便地呈现翻译中原文与译文在字/词/句层面的对应关系,帮助研究者发现原文与译文的区别性特征及这些特征与翻译中的各类转换的关系。

(三)研究思路

本研究将利用 WordSmith 对旅游文本英译文及旅游文本英文源文的词长、类符/形符比、词汇密度、词频等项目进行统计和分析,并通过对比研究的方法,探讨旅游文本英译文与原创英文旅游文本在词汇层面上各自的典型特征,以及这些特征与文本难度、用词变化、信息量、叙事态度/叙事视角及主题表达的关系;同时基于上述分析,结合 ParaConc 对旅游文本汉英平行语料进行考察,探讨特定旅游词汇的汉英翻译,以期深化学界对英译旅游文本词汇特征的认识,并对汉英旅游翻译提供有益的指导和借鉴。

三、旅游文本英译文的词汇特征及汉英旅游翻译

(一)CTETT 和 COETT 的总体语言特征统计

把“旅游文本英译文语料”(CTETT)和“旅游文本英文源文语料”(COETT)分别用 WordSmith 的 WordList 功能进行处理,得到如下统计结果:

表 1 旅游文本英译文语料(CTETT)总体语言特征统计

Text file	Tokens (running word) in text	Types (distinct words)	Type/token ration (TTR)	Standardised TTR	STTR basis	Mean word length	Mean sentence length (in words)
CTETT	82,509	8,136	10.08	42.15	1,000	4.65	20.48

表 2 旅游文本英文源文语料(COETT)总体语言特征统计

Text file	Tokens (running word) in text	Types (distinct words)	Type/token ration (TTR)	Standardised TTR	STTR basis	Mean word length	Mean sentence length (in words)
COETT	82,401	11,119	13.68	50.50	1,000	4.76	20.45

下面将基于以上统计数据,并结合 WordSmith 对子库二其他相关词汇项目的统计(如词汇密度、词频等)和 ParaConc 对子库一的相关检索,对旅游文本英译文的词汇特征及汉英旅游翻译相

关问题进行考察。

(二)词长与文本难度

词长在一定程度上反映了一个文本中用词的大小(“大词”或“小词”)和难易程度(“难懂”或“易

懂”)。根据 Edward B. Fry 的观点,英语文本的难度主要取决于句子的长短和句子中单词音节的多寡。如果句子长,且句中单词的音节又多,那么一般来说难度就大;反之,如果句子短而句中单词的音节又少,那么难度就小。<sup>⑨</sup>后来 Danielson & Bryan 也提出,可依据以下两个要素来判断英语文本的难度:1. 每个单位空间中的平均词数(实际上是对词长的测量);2. 句子的平均词数(实际上是对句长的测量)。<sup>⑩</sup>以上观点都表明,文本的难度除与句长关系密切外,与词长也有密切的关系。一般而言,文本中的平均词长值越高,则文本中大词的使用也越多,难度也越大。

以上表 1 和表 2 中的数据显示,CTETT 的平均词长(mean word length)为 4.65,COETT 的平均词长为 4.76,两者十分接近,前者比后者略低。再结合 CTETT 与 COETT 的平均句长的统计数据(20.48 vs. 20.45),可以看出旅游文本英译文与旅游文本英文源文的词汇难度非常接近,虽然前者比后者在用词上会略微简单一些,但两者并无明显差异。这与侯晋荣的研究结果一致<sup>⑪</sup>,不过在具体的平均词长的数值上与其研究略有差异。<sup>⑫</sup>

### (三)类符/形符比与用词变化

形符(token)是语料库语言学中的一个术语,类似于我们日常所说的“词”,一个文本中的形符,就是指该文本中所有的词。类符(type)是指不重复计算的形符。“类符/形符比”(type/token ratio, TTR)指文本中类符和形符的比例。它可以在一定程度上反映文本用词的变化和丰富程度。类符/形符比越高,文本的用词变化越大,用词也越丰富;类符/形符比越低,文本的用词变化越小,用词也越单调/重复。由于语料库的容量越大,其中的大量的功能词(如 the, a, of)会反复出现,形符会相应增加,但类符却未必会相应增加。这导致文本越长,功能词的重复次数也越多,类符/形符比就会越小,这样不同容量的语料库的类符/形符比就不具有可比性。因此在实际研究中,我们用“标准类符形符比”(standard type/token ratio, STTR)来衡量文本的词汇变化情况,通常是按每 1000 个形符来计算一次类符/形符比,再求其平均值。这种计算可由 WordSmith 自动完成。

以上表 1 和表 2 中的数据显示,CTETT 的标准类符/形符比为 42.15,COETT 的标准类符/形符比为 50.50。这表明与旅游文本英文源文相比,旅游文本英译文的词汇变化不够丰富,用词较为单调,词汇重复率偏高;而旅游文本英文源文用词变

化较为丰富,用到的不同词的词汇量要大于旅游文本英译文。这提示在汉英旅游翻译中,译者应该尽量使用不同的词汇来对相同或相似的意义进行表达,以使词汇变化更为丰富。

### (四)词汇密度与信息量

语言中的词分为两类:一类是开放类词(open-class words),也即实义词或实词(content words),指那些数目较大而且可以持续增生的词,这类词一般具有较为稳定或明晰或实在的词汇意义,如英语中的开放类词包括名词、动词、形容词和副词。另一类是封闭类词(closed-class words),也即功能词(function words),指数目有限且很少增生的词,这类词不大具备稳定、实在的词汇意义,或意义模糊,或主要起语法功能作用,如英语中的封闭类词包括介词、连词、限定词、代词、数词<sup>⑬</sup>、助动词和感叹词。<sup>⑭</sup>词汇密度(lexical density)是指一个文本中开放类词数量占总词数的比例,这个比例在很大程度上可以体现文本的信息量(information load)<sup>⑮</sup>,这是由于开放类词传递大部分具有实质性内容的信息,而封闭类词主要起语法功能的作用,较少传递实质性信息。因此文本中开放类词比例越高,则词汇密度越大,文本的信息量越丰富;反之,文本中开放类词比例越低,则词汇密度越小,文本的信息量则较少。另外,词汇密度与文本的难易度也有一定的关系,一般而言,词汇密度越高,文本难度越大;反之则文本难度越小。<sup>⑯</sup>

本研究对词汇密度的计算采用 Ure<sup>⑰</sup> 和 Stubbs<sup>⑱</sup> 提出的方法:词汇密度=(实词数/总词数)×100%。我们把经过用 TreeTagger 词性赋码的 CTETT 和 COETT 分别用 WordSmith 进行处理,统计其各类开放类词的数量及在总词数中的比例,结果如表 3:

表 3 CTETT 和 COETT 的词汇密度

语料 词类	CTETT		COETT		卡方值	P 值
	比例	频次	比例	频次		
开放类词	名词	27,908 33.8 %	28,747 34.8 %	20.6382	0.000***	
	动词	10,865 13.2 %	11,323 13.7 %	11.6278	0.000***	
	形容词	8,522 10.3 %	7,783 9.4 %	36.1027	0.000***	
	副词	2,918 3.5 %	3,846 4.7 %	134.0348	0.000***	
	合计	50,213 60.8 %	51,699 62.6 %	116.7946	0.000***	

表 3 显示,CTETT 和 COETT 的词汇密度分别为 60.8%和 62.6%,两者比较接近,前者略低。卡方检验显示,开放类词在两类语料中出现的频数

具有显著性差异。这表明一般情况下,与旅游文本英文源文相比,旅游文本英译文中开放类词使用的比例会小一些,因而信息量也会低一些。

再具体考察四种实词词类在总形符数中所占的比例,可以发现,名词、动词和副词的比例在CTETT中均比在COETT中略低,卡方检验显示其出现的频数具有显著性差异。其中名词在两类语料中所占的比例均为最高,分别为33.8%和34.8%;其次是动词,分别为13.2%和13.7%。王克非的研究显示,在英语原创文学语料中,名词所占比例为20.65%,动词所占比例为19.88%<sup>⑨</sup>,对照本研究,可发现英文旅游文本中(包括英文源文本及英译文本),名词所占比例要大大高于英语文学文本中的比例,而动词所占比例则要大大低于英语文学文本中的比例。究其原因,主要在于旅游文本区别于文学文本的一个典型特征:旅游文本的叙事方式总体上是介绍性的。一方面旅游文本会涉及大量的人名、地名及事物名的介绍(如对旅游景点/设施、风俗习惯、历史事件及器物用品等的介绍),因此名词的使用自然会非常频繁;另一方面,对事物介绍性的叙事方式会导致旅游文本中所使用的静态性的词汇(如名词)趋多,而动态性的词汇(如动词)趋少,因此名词在英语旅游源文本及英译文本中的比例大大高于英语文学文本中的比例,而相应的动词比例则大大低于英语文学文本中的比例。

考察表3中形容词出现的比例,发现在CTETT中形容词出现的比例比在COETT中高将近0.9%,究其原因,可能在于“英文旅游文体大多风格简约……行文用字简洁明了,表达直观通俗,注重信息的准确性和语言的实用性”,而汉语旅游文本多华丽的渲染与铺陈,“文采浓郁”且“大多仰仗辞藻的渲染”<sup>⑩</sup>,而在汉英旅游翻译中,这些“渲染铺陈、文采浓郁”的表述经常会译为形容词,从而导致旅游文本英译文中形容词的比例高于旅游文本英文源文中的比例。这提示在汉英旅游翻译中,对“辞藻华丽”的汉语原文,我们可采用简略化或淡化的处理方法,限制形容词的过度使用,使译文风格简约、直观通达,这样译文才更符合英文旅游文本的词汇规范和行文风格。

#### (五)人称代词与叙事态度/叙事视角

把CTETT和COETT分别导入WordSmith中,统计两类语料中人称主格代词的使用情况,得到如下统计结果:

表4 CTETT和COETT中主格代词的使用统计

语料	CTETT	COETT
总形符数	82509	82401
第一人称主格代词(I, we)频次及比例	65 0.079%	305 0.37%
第二人称主格代词(you)频次及比例	232 0.28%	802 0.97%
第三人称主格代词 (he, she, it, they)频次及比例	1017 1.2%	760 0.92%

考察表4中三类人称主格代词在CTETT和COETT中的使用情况可以发现,第一人称主格代词(I, we)在COETT中的比例(0.37%)大大高于在CTETT中的比例(0.079%),可见第一人称主格代词在原创英文旅游文本中的使用比例要大大高于旅游文本英译文中的比例。从语用学的角度讲,“在特定情况下说话人对指示方式的特定选择表明了说话人对听话人的态度、情感和相互关系”<sup>⑪</sup>,而“指示方式”,首先就涉及人称指示,即人称代词的使用。我们知道,第一人称代词的使用可以拉近说话人/作者和听话人/读者之间的距离,产生一种亲近感或亲切感。本研究显示,与旅游文本英文源文相比,旅游文本英文源文会更多地使用第一人称代词,从而拉近说话人/作者(通常为旅游公司)和游客之间的距离,以便吸引更多的游客来观光旅游,达到旅游推介的目的。而汉语旅游文本中,则相对来说较少使用第一人称代词来达到以上效果。这提示我们在汉英旅游文本翻译中,可能需要适当地进行人称代词的转换,更多地使用第一人称代词,从而拉近说话人/作者与游客之间的距离,吸引更多的外国游客来参观旅游,以便更好地实现汉英旅游翻译的目的。例如:

九龙酒店秉承个性化的温馨服务理念,拥有严格专业的管理团队。饭店所有员工以热情诚挚的服务殷切期待您的光临。

(1) Jiulong Hotel embraces personalized and warm service idea, and boasts strict and professional management team. The staff of the Hotel are ready to welcome you with cordial hearts.

(2) With customer-oriented service and professional management team, we promise to offer you an unforgettable stay in our hotel.

上例选取自本研究的平行语料库。其中第一个译文把原文的“饭店所有员工”直译为“The staff

of Jiulong Hotel”,虽忠实于原文意义,但在英语读者看起来,显得很有距离感,缺乏足够的感召功能(vocative function)。第二个译文把“饭店所有员工”进行人称转换,译为第一人称主格代词“we”,拉近了酒店和游客之间的距离,具有更强的感召功能。

再来看看第二人称主格代词(you)在 CTETT 和 COETT 中的使用情况。CTETT 中 you 的使用比例为 0.28%,在 COETT 中为 0.97%,可见第二人称主格代词 you 在原创英文旅游文本中的使用比例也大大高于旅游文本英译文中的比例。究其原因,这与第一人称主格代词“I / we”在 CTETT 和 COETT 中的使用情况类似。汉英旅游翻译中,可能需要适当地进行人称代词的转换,更多地使用第二人称代词。

#### (六)词频与主题表达及特定旅游词汇的翻译

把经过 TreeTagger 词性赋码的 CTETT 和 COETT 分别导入 WordSmith 中,载入停用词表(stoplist)和词形还原表(lemma list),然后统计两类语料中出现频次最高的居前 15 位的名词、动词和形容词,得出其统计结果。下面分别对这三类词的词频及其特性讨论如下。

##### 1. 名词词频

表 5 为 CTETT 与 COETT 中出现频次最高的前 15 个名词的统计结果。

表 5 CTETT 与 COETT 中频次最高的前 15 个名词

语料	CTETT			COETT		
	排序	词项	词频	排序	词项	词频
名词	1	river	332	1	city	311
	2	mount(ain)	319	2	park	290
	3	city	271	3	hotel	243
	4	tower	68	4	beach	192
	5	dynasty	259	5	world	183
	6	area	238	6	street	159
	7	year	229	7	day	158
	8	gorge	217	8	time	156
	9	China	192	9	museum	153
	10	pavilion	184	10	home	151
	11	people	184	11	island	144
	12	temple	183	12	place	141
	13	stone	167	13	art	136
	14	water	165	14	state	130
	15	world	152	15	area	128

从表 5 可看出,CTETT 与 COETT 中出现的高频名词具有不同的特点:CTETT 中,排在第 1、2、4 位的词分别是 river, mountain/mount 和 tower;而 COETT 中,排在第 1、2、3、6 位的词分别是

city, park, hotel 和 street,这说明旅游文本英译文中对自然(及人文)景观特别是山川河流的介绍占重要分量,而相对而言旅游文本英文源文对城市景观及旅游设施的介绍占很大比重。CTETT 中, dynasty, year 和 people 分别排在第 5、7 和 11 位,而表达类似概念的词在 COETT 中没有出现在前 15 位,这表明旅游文本英译文倾向于通过介绍旅游景点的悠久历史和人文故事来吸引游客,而旅游文本英文源文则不太注重这一方面。CTETT 中,体现中华人文建筑的词 tower, pavilion, temple 分别列第 4、10、12 位,而这些词均未见于上表 COETT 中——与此形成对照的是,在 COETT 中, museum 一词出现了 153 次,排在第 9 位, art 一词出现 136 次,排在第 13 位;而在 CTETT 中, museum 一词仅出现 33 次, art 仅出现 21 次,其排位都在 60 位之后,这表明旅游文本英译文更强调对具有中华建筑特色的“楼/塔”、“亭”和“寺/庙”等进行介绍,而旅游文本英文源文则更重视对博物馆等城市人文艺术建筑进行介绍。

值得注意的是, home 一词在 COETT 中出现了 151 次,排第 10 位,而在 CTETT 中仅出现了 34 次,排在 60 位之后。为何该词在两类语料中出现的频率相差如此之大?我们利用 WordSmith 的 Concord 工具考察 home 在 COETT 中的语义及搭配情况,发现在原创英文旅游文本中, home 除了表示常见的 place where one lives, esp with one's family(家)的意思外,还经常用来表示:(1) place where an animal or a plant is native or most common; habitat ([动植物的]栖息地, 栖息处, 生息地);(2) place from which sth originates(发源地, 发祥地)。在表示这两个意思时, home 之后与介词的常见搭配为 home to 或者 home of, 如图 2、图 3 所示。

Home 一词出现在 COETT 中共计 151 次,其中有 70 次该词表示“(动植物的)栖息地, 栖息处, 生息地”(55 次)或“发源地, 发祥地”(15 次)之义,占该词在 COETT 中出现频数的近二分之一。而在 CTETT 中, home 出现 33 次,其中只有 1 次表示“发源地, 发祥地”之义,仅占三十三分之一。这提示在汉英旅游翻译中,在涉及“栖息地, 栖息处, 生息地”或者“发源地, 发祥地”这些表述时,译者的翻译有需要改进之处,特别是译者应注意在此语境中 home 一词的灵活使用。我们用 ParaConc 分别导入 CCTT 和 CTETT,用“栖息”作为关键词进行检索,发现在翻译中的确有些涉及“栖息”这一表述的汉语原文没有得到妥善处理。例如下面两例:

Concord										
File Edit View Compute Settings Windows Help										
N	Concordance	Set	Tag	Word #	Sen	Sen	Par	Par	Hea	Hea
2	, the Gálapagos Islands are also home to an incredibly diverse collection			42,85	2,2	31	0	53		
3	Avenue Historic Shopping District (also home to an eclectic collection of locally			34,29	1,6	48	0	42		
4	your hand at real cowboy stuff. It's also home to one of the world's greatest			29,42	1,5	27	0	36		
5	worth a visit. The Rialto Market is also home to the ancient Church of San			46,63	2,4	47	0	57		
6	artists, the neighborhood is also home to the perennially popular Beach			65,43	3,2	50	0	81		
7	views of the Manhattan skyline. It's also home to Montague Street Bagels, a			62,44	3,1	17	0	77		
8	located in the heart of the city. It's also home to Anish Kapoor's Cloud Gate			76,65	3,7	9%	0	94		
9	in the middle of wine country. It's also home to a small bunch of breweries			19,29	1,0	19	0	24		
10	games: Lakefront Brewery. It is also home to mascot Bernie Brewer's			20,22	1,0	33	0	25		
11	Pavilion (designed by Frank Gehry and home to the Grant Park Music			76,67	3,7	47	0	95		
12	jewel loaded with Lalique glass and home to some fine modern Italian and			79,97	3,5	71	0	99		
13	the Collegiate Peaks Wilderness Area, home to eight of the state's fourteeners			27,50	1,4	88	0	34		
14	, it's a good idea for people back home to know where you are. Whether			3,884	19	82	0	5%		
15	Museum and the Parc des Bastions -- home to the oldest buildings of the			68,82	3,4	85	0	85		
16	Winter Olympics, also happens to be home to Utah's only Alpine coaster.			40,24	2,1	77	0	50		
17	Hall (H?tel-de-Ville), which has been home to the government of the			68,44	3,3	62	0	84		
18	up to the grand family name it carries. Home to an art center, the GE			64,48	3,2	8%	0	79		
19	bridges in film, TV and art. Chinatown Home to the largest population of			64,08	3,2	12	0	79		
20	beer tour needs to start in Fort Collins. Home to Belgian-inspired Funkwerks			19,76	1,0	5%	0	24		
21	and markets. Madison Square Garden Home to approximately 320 events			64,39	3,2	23	0	79		
22	Alps, and at the top is Stelvio Glacier, home to some of the continent's best			32,60	1,7	67	0	40		
23	(which celebrates 100 years in 2012) is home to the Boston Red Sox, a team			74,51	3,6	29	0	92		
24	in a unique way. This boutique resort is home to one-bedroom Guesthouses			54,16	2,7	40	0	67		
25	Monica Pier The Santa Monica Pier is home to the world's only solar-powered			80,78	3,5	38	0	10		
26	ground at Fifth and Arch Streets is home to the graves of such historic			70,61	3,4	53	0	87		
27	it just so happens that Angkor Wat is home to one of the best sunrises on			14,18	78	79	0	17		
28	southern tip of Manhattan, this park is home to The Sphere -- a large,			63,12	3,1	53	0	78		
concordance collocates plot patterns clusters timeline filenames source text notes										
57 entries	Row 1									new World War II airplanes Home to one of the world's larges

图2 home+to 搭配检索结果

Concord										
File Edit View Compute Settings Windows Help										
N	Concordance	Set	Tag	Word #	Sen	Sen	Par	Par	Hea	Hea
1	for. Inform family and friends back home of your whereabouts and			5,316	29	44	0	7%		
2	San Francisco's iconic Ferry Building, home of the famous Ferry Plaza			65,16	3,2	48	0	80		
3	cred extends to Chase Field, home of the Arizona Diamondbacks. It			33,33	1,6	79	0	41		
4	. Walk around leafy Brooklyn Heights -- home of gorgeous brownstone			62,42	3,1	30	0	77		
5	, the Dolby Theatre is the official home of the Academy Awards, and			80,71	3,5	58	0	99		
6	to Hawaii Volcanoes National Park, home of two active volcanoes: Kilauea			39,38	2,0	69	0	49		
7	, you may stumble on to Wall Street, home of the heavily guarded New York			61,96	3,1	42	0	76		
8	downtown Phoenix and nearby Tempe, home of Arizona State University.			33,74	1,6	82	0	42		
9	and how to best enjoy your time at the "Home of the Beachfront Li'l Grass			58,67	2,5	81	0	72		
10	if you stay on it. A journey to the home of the wizard of Oz There's no			13,58	75	33	0	17		
11	from the Miller signs and head to the home of the other beer sold at			20,20	1,0	71	0	25		
12	certainly qualifies. Most notably the home of Old Faithful Geyser, the park			39,85	2,1	23	0	49		
13	Berlin. The Berlin area was once the home of major peach orchards, and the			23,23	1,2	40	0	29		
14	or head back to the Ferry Building, the home of Slanted Door, one of the			65,51	3,2	67	0	81		
15	on every block. Tee Off at The western home of Golf in America An ideal stop			51,22	2,6	16	0	63		
16	Weiskopf. Known as "The Western Home of Golf in America," the resort			51,26	2,6	29	0	63		
17	Jones, Sunnylands was the winter home of Ambassador Walter			51,34	2,6	50	0	63		
concordance collocates plot patterns clusters timeline filenames source text notes										
17 entries	Row 1									rm family and friends back home of your whereabouts and situ

图3 home+of 搭配检索结果

1. 黄鹤楼脚下的黄鹤矶,因附近江洲有天鹅栖息而得名;蛇山形似一条伸入江中饮水的长蛇;龟山则像浮在水上的巨鼉。

——《天下绝景黄鹤楼》

The Huanghuji at the foot of the Yellow Crane Tower was named so for the fact that nearby there was a dwelling area for swans. The Snake Hill zigzags its way forward, like a long snake drinking water across the River, while the Tortoise Hill looks like a giant Yuan (a kind of soft-shelled turtle) floating on the water.

—*The Superb Yellow Crane Tower under Heaven*

2. 此楼脚下的山矶名“黄鹤矶”,江湾名

“黄鹤湾”,古时,这一带天鹅繁衍栖息……

——《天下绝景黄鹤楼》

There was a rock called “Huang-He-Ji” (yellow crane rock) protruding over the river at the foot of the tower, and the river bend was hence called “Huang-Hu-Wan” (yellow swan bay) where flocks of swans dwelled in the ancient times. ...

—*The Superb Yellow Crane Tower under Heaven*

第一个例子中,“因附近江洲有天鹅栖息而得名”显然译得非常生硬,可改译为“... was named so as its nearby rivers and islets used to be the home to swans (‘swan’ is pronounced ‘huanghu’ and ‘islet’ is pronounced ‘ji’ in Chinese pinyin)”。第二个例子中,“这一带天鹅繁衍栖息”译得

也不地道,可改译为“which was the home to swans in ancient times.”

2. 动词词频

表 6 为 CTETT 与 COETT 中出现频次最高的前 15 个(实义)动词的统计结果。

表 6 CTETT 与 COETT 中  
频次最高的前 15 个动词

语料	CTETT			COETT		
	排序	词项	词频	排序	词项	词频
动词	1	build	160	1	take	250
	2	make	151	2	get	184
	3	locate	135	3	make	148
	4	know	112	4	go	136
	5	call	109	5	offer	133
	6	stand	98	6	know	131
	7	take	94	7	see	89
	8	go	91	8	visit	86
	9	cover	82	9	include	85
	10	look	80	10	want	80
	11	come	79	11	say	79
	12	see	76	12	enjoy	77
	13	become	71	13	find	75
	14	say	65	14	look	74
	15	visit	57	15	walk	71

在(实义)动词的使用上,CTETT 和 COETT 呈现出不同的特点。首先我们来考察一下行为动词的词频。在 COETT 中,词频排在前四位的词分别为 take, get, make, go, 这些词均为英语中最常见、也是最常用的共核(common core)行为动词,在旅游文本英文源文中它们也得到了广泛的运用,词频依次排在表中的最前列。不过在 CTETT 中,仅有 make 一词出现在前四的位置里(排第 2 位),take 和 go 分别排在第 7 和 8,而 get 没有出现在前 15 位中。我们知道,地道的英语表述离不开对英文中的共核词汇、特别是那些最常见的共核动词(如 take, get, make, go)的灵活运用,而在汉英旅游翻译中,显然这种对最常见的共核动词的灵活运用还做得不够。

表 6 中 offer 这一动词的排位情况颇值得关注。在 COETT 中该词共出现 133 次,高居词频表中的第 5 位;而在 CTETT 中,该词仅出现 9 次,排在非常靠后的位置。考察 offer 在 COETT 中的搭配情况,发现该词经常用于英文旅游风景/景点“呈现状态”的描述中,而在 CTETT 中该词较少用于这一语境。出现这一状况的原因,可能在于旅游风景或旅游景点“呈现状态”的描述中,汉语中常常不太容易找到一个与 offer 直接对应的词(ready-

equivalent),如下例:

The rooftop pool offers some of the best views in the city, and the restaurant is legendary.

—Old San Juan: The Best Quick (and Cheap!) Weekend Getaway

这句话如果译成汉语,可译为:

从天顶泳池可以看到城市最美的风景,而且这里的餐厅也享誉卓著。

这里 offer 一词并未译为“提供”、“呈现”之类,而是转换为“可以看到”,而这个表述并非 offer 的直接对应的词。相应地,把汉语的“(从……)可以看到……”这个表述译为“… offer …”也不容易想到。由此可看出,在旅游风景的描述中,汉语中有时不大容易找到一个贴切的词与 offer 直接对应,而这会导致翻译转换中的一些困难。这提示我们在汉英旅游翻译中,译者需要灵活处理原文,在翻译旅游风景/景点“呈现状态”的描述时,要特别注意 offer 一词的灵活运用。我们用 ParaConc 导入 CCTT 和 CTETT,用“offer”作为关键词进行检索,发现在有些句子的翻译中,译者对原文进行了灵活处理,offer 一词运用得当,译文地道自然。例如下例:

云南有 26 个少数民族,是中国少数民族种类最多的省份。各民族的服饰、建筑、风俗、歌舞、饮食等,形成了一幅美丽的风情画卷。

——《云南印象》

Home to 26 ethnic groups—the largest number in China—Yunnan Province offers tourists a cultural feast of unique ethnic costumes, architecture, cuisine, songs and dances, and rituals.

—A Profile of Yunan

本例中,“形成了”这一表述用“offers”来译,不仅意义忠实于原文,且简洁地道。虽然“形成了”也可译为“forms”,但与该词相比,offers 一词更为贴切,因为 offers 能够更好地体现旅游地域(Yunnan Province)对游客的“给予”之意(offers tourists sth; offers sth. to tourists),从而能更好地实现旅游翻译的呼唤功能,而 forms 则很难体现这种意义、实现这种功能。

总之,旅游文本具有强烈的呼唤功能。“英语源语旅游文本在劝导人们时用词平实明快、简洁明了,表达直观通俗,呼唤游客去行动,去思考,去感受。”<sup>②</sup>英语旅游文本在用词特别是动词使用上的以上特征,值得我们注意,并在汉英旅游翻译中加

以借鉴。

### 3. 形容词词频

表 7 为 CTETT 与 COETT 中出现频次最高的前 15 个形容词的统计结果。

表 7 CTETT 与 COETT 中频次最高的前 15 个形容词

语料	CTETT			COETT		
	排序	词项	词频	排序	词项	词频
形容词	1	ancient	199	1	best	139
	2	Chinese	186	2	great	127
	3	scenic	157	3	local	91
	4	famous	143	4	old	121
	5	long	136	5	national	107
	6	yellow	109	6	long	85
	7	high	107	7	big	78
	8	cultural	103	8	historic	73
	9	great	98	9	small	65
	10	natural	92	10	large	63
	11	beautiful	91	11	new	60
	12	white	88	12	hot	60
	13	national	87	13	good	57
	14	historical	83	14	high	57
	15	large	76	15	famous	55

表 7 显示,CTETT 中,排第 1、2 和第 8 位的形容词分别是 ancient、Chinese 和 cultural,这些词均不见于上表 COETT 中,这表明与原创英语旅游文本相比,旅游文本英译文多强调景点的历史和文化元素,以此吸引游客。COETT 中排第一位的词是 best,排第二位的是 great,表明旅游文本英文源文中,倾向于强调某一景点或旅游设施为同类中“最好”、“最佳”以及“优秀”、“杰出”这些特点来吸引游客。

令人尤为感兴趣的是 scenic 一词的使用情况。在 CTETT 中,scenic 一词共出现 157 次,排第 3 位;而在 COETT 中,该词仅出现 15 次,排位相当靠后。CTETT 和 COETT 均为旅游文本语料,其中肯定都不乏对“风景”、“景点”、“景区”的表述和描绘,照理说两类语料都应该会频繁地用到 scenic 一词,可为何该词在这两类语料中出现的频率相差这么大?

利用 WordSmith 的 Concord 工具考察 CTETT 中 scenic 的搭配情况,发现 scenic 最常见的搭配为 scenic spot(s),共出现 85 次,其他的搭配为 scenic area(s) (39 次),scenic region(s) (13 次),scenic attraction(s) (6 次),scenic park(s) (4 次),等等。显然这些搭配是用来翻译原文的“景点”、“景区”、“风景区”之类的表述的。问题是这些搭配是否也经常出现在原创英文旅游文本中呢?

考察 COETT 中 scenic 的搭配情况,发现 scenic 出现的 15 次中,其搭配如下:

scenic drive (2); scenic route (2); scenic byway (2); scenic road (2); scenic (hiking) trails (1); scenic vistas (1); scenic (French) island (1); scenic landscapes (1); scenic expanses of .... (1); is (so) scenic ... (2)。

值得注意的是,以上 scenic 在 COETT 中呈现的那些搭配,如 scenic drive, scenic route, scenic byway, scenic road 等,在 CTETT 中一次都没有出现;而在 CTETT 中 scenic 呈现的某些高频次的搭配,特别是 scenic spot(s) 这一表述,在 COETT 中也一次都没出现。这一结果令人感到有些意外,它使人想到这样一些问题:我们在汉英旅游翻译中,对 scenic 一词的使用是否妥当?我们理所当然地认为“景点”应译为 scenic spot(s),<sup>③</sup>不过这个译文是否足够地道和恰当?或许 scenic spot(s) 并非“景点”理想的译文?

具体考察以上 COETT 中 scenic 的搭配情况,可以看出 scenic 的语义韵(semantic prosody)呈现这样一种特点,即该词似乎更多地与表示“(沿着)线路、路径”的名词搭配,而较少与表示“某一特定地点”的名词搭配。可以设想,或许正是这一原因,导致英文中 scenic 一词较少与 spot 或 site 等名词搭配。

为验证以上假设,我们把 scenic spot(s) 这一搭配输入英语国家语料库 (British National Corpus, BNC) <sup>④</sup>,结果在这个库容为一亿词的权威的英语语料库中,scenic spot(s) 的搜索结果居然为零。这也提示 scenic spot(s) 这一表述在英语中可能使用得相当有限,至少其使用并不像我们想象中那么普遍。

具体考察 scenic 在 BNC 中的搭配情况,发现该词常见的搭配有:

scenic route; scenic drive; scenic railways; scenic promenades; scenic run; scenic road; scenic journey; scenic flight; scenic canter; scenic walk; scenic coastline

从以上搭配可看出,scenic 一词的确更多地与表示“(沿着)线路、路径”的名词搭配,而较少与表示“某一特定地点”的名词(如 spot) 搭配——不过,检索也发现,scenic 倒是可以与表示某一“地域/区域”的名词的搭配,这种搭配主要出现于 scenic area 这一表述中(相当于“[风]景区”)。

那么 spot 一词前可与什么词搭配表示“景点”呢?



在 *Oxford Collocations Dictionary for Learners of English* 中考察 spot 前与其他词语的搭配情况,发现 spot 前可与 beauty 搭配,如: The lake is one of the local beauty spots. 而根据 *Oxford Advanced Learner's Dictionary (OALD)*, beauty spot 意为 a place of natural beauty; *Collins Advanced Learner's Dictionary* (5<sup>th</sup> edition) 对 beauty spot 的解释为: a place ... that is popular because of its beautiful scenery. 这与“景点”的意思比较吻合。我们以 beauty spot(s) 为关键词在 BNC 中进行检索,可以搜索到 96 个检索结果,如:

Guests can enjoy trips to local beauty spots, dancing at a local hotel, and visits to local theatres.

而在 COETT 中,也可以检索到两个含有 beauty spot 的句子。可见,spot 可与 beauty 搭配(即 beauty spot)表示“景点”之意。<sup>②</sup>

根据 *Oxford Collocations Dictionary for Learners of English*, spot 前还可与 tourist 搭配。在 *Collins Advanced Learner's Dictionary* (5<sup>th</sup> edition) 中可以找到下面的例句:

They stayed at several of the island's top tourist spots ...

又如在 *Macmillan English Dictionary* 中可以找到如下例句:

... one of the region's best-known tourist spots

以 tourist spot(s) 为关键词在 BNC 中进行检索,共搜索到 14 个结果。如:

It was an almost empty theatre in Long Beach, Long Island, a tourist spot with a boardwalk and beautiful beaches.

可见 tourist spot 也可用来表示“景点”之意。

另外,travel spot 在本研究的 COETT 也可以搜索到,如:

Top Family Travel Spots on Maryland's Eastern Shore

可见 beauty spot, tourist spot 及 travel spot 都可作为“景点”的英译文,而 scenic area 可作为“(风)景区”的英译文。至于 scenic spot 这一译文,虽可以为英语国家读者所理解(抑或可用于口语化的表述中),但并非英文旅游文本中正式、地道和常用的表述。

#### 四、小结

本研究基于汉英旅游文本平行语料库和英文

旅游文本可比语料库,对旅游文本英译文的词汇特征及特定旅游词汇的翻译问题进行了探讨,重点考察了旅游文本英译文的词长与文本难度、类符/形符比与用词变化、词汇密度与信息量、人称代词与叙事态度/叙事视角,以及词频(名词词频、动词词频和形容词词频)与主题表达及特定旅游词汇的翻译问题。研究发现,1. 旅游文本英译文与旅游文本英文源文的词汇难度非常接近,只是前者比后者略低,但并无明显差异。2. 与旅游文本英文源文相比,旅游文本英译文的词汇变化不够丰富,词汇重复率偏高。在汉英旅游翻译中,译者应尽量使用不同的词汇来对相同或相似的意义进行表达,以提高词汇变化的丰富程度。3. 与旅游文本英文源文相比,旅游文本英译文的词汇密度略低,其信息量也略低一些。4. 与旅游文本英文源文相比,旅游文本英译文中第一人称和第二人称(主格)代词的使用比例明显偏低,故在汉英旅游翻译中,需要适当地进行人称代词的转换,更多地使用第一人称和第二人称代词,从而拉近旅游文本与游客之间的距离,更好地实现汉英旅游翻译的目的。5. 旅游文本英译文中对自然景观特别是山川河流的介绍占重要分量,而旅游文本英文源文中对城市景观及旅游设施的介绍占较大比重。6. 旅游文本英译文中,对某些最常见的共核行为动词(如 take, get)的灵活运用做得不够,对此需加以改进。7. 汉英旅游翻译中,要特别注意对名词“home”(表“栖息地”及“发源地”之意)及动词“offer”(表“[景色]呈现状态”之意)的灵活应用。另外,形容词“scenic”与名词“spot”一词搭配构成的短语“scenic spot”并非“景点”一词理想的译文,“tourist spot”、“travel spot”及“beauty spot”比“scenic spot”更适合作为“景点”一词的对应译语。

#### 注释

①陈刚:《旅游翻译与涉外导游》,北京:中国对外翻译出版公司,2004年。

②方梦之、毛忠明:《英汉-汉英应用翻译教程》,上海:上海外语教育出版社,2004年,第274-318页。

③④贾文波:《应用翻译功能论》,北京:中国对外翻译出版公司,2004年,第106-156页;第109-111页。

⑤伍锋、何庆机:《应用文体翻译:理论与实践》,杭州:浙江大学出版社,2008年,第317-352页。

⑥李德超、王克非:《新型双语旅游语料库的研制和应用》,《现代外语》2010年第1期。

⑦⑧侯晋荣:《基于语料库的旅游文本语言特征及语篇分析》,《菏泽学院学报》2011年第6期。

⑨肖庚生、陈欣:《旅游汉英双语平行语料库的建设与

应用》,《湖南科技学院学报》2012年第10期。

⑧⑨张苇、韩江洪、夏蓓洁:《基于自建语料库的旅游英语文体特征类比研究》,《合肥工业大学学报(社会科学版)》2015年第2期。

⑩Fry, E. B. *Elementary Reading Instruction*. New York: McGraw-Hill, 1977.

⑪李燕、张英伟:《〈博雅汉语〉教材语料难度的定量分析——兼谈影响教材语言难度的因素和题材的选择》,《云南师范大学学报(对外汉语教学与研究版)》2010年第1期。

⑫根据侯晋荣的研究,汉语旅游文本英译文的平均词长为4.71,参照文本(即英语旅游文本原文)为4.81。

⑬数词虽然可以无限增生,但其意义仅限于数字的表达,因此也归为封闭类词。

⑭章振邦:《新编英语语法教程》,上海:上海外语教育出版社,1983年,第9-10页。

⑮⑯Baker, M. "Corpora in Translation Studies: An Overview and Some Suggestions for Future Research." *Target*, 2 (1995): 223-243; 237-238.

⑰Ure, J. "Lexical Density and Register Differentiation." In Perren, G., and J. L. M. Trim, eds. *Applications of Linguistics*, 443-452. Cambridge: Cambridge University Press, 1971.

⑱Stubbs, M. "Lexical Density: A Technique and Some Findings." In Coulthard, M., ed. *Talking about Text: Studies Presented to David Brazil on His Retirement, Discourse Analysis Monograph*, 27-42. Birmingham: ELR, 1986.

⑲王克非:《语料库翻译学探索》,上海:上海交通大学出版社,2012,第63页。

⑳何兆熊:《新编语用学概要》,上海:上海外语教育出版社,2000年,第66页。

㉑国内权威的汉英词典,如《新时代汉英大辞典》(吴景荣、程镇球,2000)、《汉英词典》(姚小平,2010)、《新世纪汉英大辞典》(惠宇,2004)、“景点”一词都毫无例外地被译为“scenic spot”或“scenic site”。

㉒BNC主页网址: <http://www.natcorp.ox.ac.uk/>, 检索时间为2015年12月18日。

㉓beauty spot 还有另外一个意思,即 a small dark mark on a woman's face, used when you think it is attractive (OALD) (美人痣)。

责任编辑 王雪松

## A Corpus-based Study on Lexical Features of C-E Translations of Tourism Texts and C-E Tourism Translation

Xiong Bing

(School of Foreign Languages, Central China Normal University, Wuhan 430079)

**Abstract:** The present thesis, based on the Chinese-English (C-E) parallel corpus of tourism and English comparable corpus of tourism, makes an exploration of the lexical features of C-E translations of tourism texts and C-E tourism translation. The study finds that, translated English tourism texts (TETT) and original English tourism texts (OETT) are very similar in their lexical difficulty, with the former's lexical difficulty being slightly lower than the latter's, though there is no significant difference between them. Compared with OETT, TETT has a lower lexical variation and higher lexical repetition, and exhibits a lower lexical density and accordingly has a lower information load. The first and second person pronouns are used far less frequently in TETT than in OETT. In TETT, the introduction to natural landscapes (esp. mountains and rivers) accounts for a major part as compared with the introduction to city sights and tourism facilities which takes up an important part in OETT. In C-E tourism translation, the translator is supposed to pay special attention to the flexible use of common core verbs (e. g. take, get) as well as such nouns as “home” and such verbs as “offer.” Besides, the expression “scenic spot” should not be regarded as an ideal equivalent for “景点”—“tourist spot”, “travel spot” or “beauty spot” may be better equivalent for “景点” than “scenic spot.”

**Key words:** corpus; English translated tourism text; lexical features; translation