

基于语料库的人机文学译本语言特征对比研究

——以《傲慢与偏见》三个译本为例

韩红建^{1,2}, 蒋 跃¹

(1. 西安交通大学 外国语学院 陕西 西安 710049; 2. 西安交通大学城市学院 外语系 陕西 西安 710018)

摘要: 本研究基于由《傲慢与偏见》两个人工译本与一个在线机译本及其英语原文构成的平行语料库,从词法和句法层面对比了人机译本的16个语言计量特征及风格特征。结果表明,在线译本在词法、句法和情感表达等方面都远不如人工译本,在代词、助词和连词使用上“显化”明显,在标准类型比和词汇密度方面“简化”不足,而句法层面有“明晰化”倾向。这与其在词性转换、上下文语境判断、情感识别与时态标记识别等方面能力不足有关。两个人工译本在词汇和句法层面也呈现明显的风格差异,其“显化”、“简化”和句法“明晰化”的程度也各有不同,其成因与不同的翻译目的和译者主体性有关。

关键词: 语料库; 语言计量特征; 人工译本; 在线译本

中图分类号: H059 **文献标识码:** A **文章编号:** 1000-5544(2016)05-0102-05

Abstract: Based on a parallel-corpus of *Pride and Prejudice* consisting of two human translations, one online-machine translation and the original text, this study compares the 16 quantitative and stylistic features of the human translations with those of the online translation. Results of the study indicate that the online translation is far behind human translations at lexicon and syntactic levels and in sentiment expression. It is abundant in features of “explicitation” such as using pronouns, auxiliaries and connections but it lacks “simplicity” in standard type/token ratio and lexical density, and fails to show the tendency of syntactic “explicitation”. These may be attributed to its weakness in identification and transformation of parts of speech and identification of the contextual meanings, semantic sentiments and tense of the original sentences. The two human translations show different lexical and syntactic features and differ in degree of “explicitation”, “simplification” and syntactic “explicitation”, which are probably due to the different translation purposes and subjectivities of the translators.

Key words: corpus; quantitative features; human translation; online translation

DOI:10.16362/j.cnki.cn61-1023/h.2016.05.023

1. 引言

基于网络的在线机器翻译(简称“在线机译”)从基于语料库的机器翻译(corpus-based machine translation)中衍生而来,属于机器翻译的一种。目前互联网上翻译质量较高、用户数量大的在线翻译工具主要有百度、谷歌和有道等。在线机译通过汇集海量的双语网页及文档,以整句为单位,使用统计算法对原文进行多重模糊匹配,并结合语法规则进行优化、校正和消歧后得到翻译结果,已成为众多机器翻译系统开发者的首选(冯志伟 2010: 28)。在线机器翻译是自然语言处理的一个分支,与计算语言学、自然语言理解之间有着密不可分的关系(崔磊、周明 2014: 31)。它融合了统计和深度学习方法,借助海量计算机模拟的神经元,模仿人脑“理解语言,生成译文”。然而,由于规则覆盖面有限和语料库数据的稀疏,其生成的译文准确度低,可读性较差,也不乏“愚蠢”的错误(Tucker 1987: 86)。提高在线机译的质量充满挑战,被列为21世纪十大科技难题之首。有专

家指出,要提高机器翻译的质量,首先要解决的是语言本身问题而不是程序设计问题(李浩、金文 2011: 12)。

研究在线机译与人工翻译在语言特征上的共性和差异,既是从语言本身出发,探究提高机译质量的举措,也是对翻译共性和译者风格研究的拓展。在当今“互联网+”的时代,在线机译在科技、法律和商务翻译等领域发挥着重要作用,但在注重文化和情感传递的文学翻译领域是否具有适用性,却有待探讨。计算机语料库的出现为我们从海量文本中提取语言计量特征,探索语言使用规律提供了技术支撑。同为翻译,在线机译和人工翻译的译本在语言特征上肯定会表现出一致性或区别性的特征,即翻译的特征及翻译的共性(王克非、胡显耀 2008: 19)。随着各类译学语料库在全国各个不同的高校的建成,翻译的共性研究开始了大讨论时期(张律、胡东平 2012: 114)。王克非(2012)、胡开宝(2011)、黄立波(2007)和肖忠华(2012)等学者都曾对翻译汉语的语言特征进行了深入研究。

然而,采用语料库技术和语料库翻译学研究范式,

对在线机译与人工翻译的文学译本进行语言特征对比研究的文献并不多。蒋跃(2014: 98-102)以及蒋跃、董贺(2015: 69-81)开展了基于语料库的人工译本与在线机器译本的语言计量特征对比研究。这些研究为我们研究译本语言特征提供了研究依据和重要借鉴。但他们的研究由于语料库规模偏小及考察的语言特征范围有限,且对翻译中占有重要位置的文学翻译并未进行专门研究,未免有一些缺憾。

鉴于此,本研究自建了由简·奥斯汀《傲慢与偏见》的王科一、孙致礼以及百度在线汉译本与英语原著构成的一对三平行语料库。应用语料库检索工具对英语原著与三个译本在词汇和句法层面的数据进行了提取统计和对比研究,并与兰开斯特现代汉语语料库(LCMC)、浙江大学汉语译文语料库(ZCTC)和汉英对应语料库(CEPC)的相关数据进行对比分析。本文拟研究如下问题:1)在线机译译本与人工译本各呈现出哪些语言计量特征和风格特征?其成因何在?2)两种译本与原创汉语语言相比,体现了翻译语言的哪些共性特征?3)人机译本语言特征的差异对改进机译质量有何启发?

2. 语料选取与方法

2.1 语料选取与加工

在自建的语料库中,小说的英文版为牛津大学出版社1970年版本。中译本采用王科一译本(1955年出版)、孙致礼译本(1990年出版)和百度在线译本。王科一译本乃《傲慢与偏见》第一个中译本,以语言准确、通顺而被赞为“一部较好的译本”(孙致礼2009: 42)。孙致礼译本(简称“孙译本”),因语言“自然流畅、准确生动”(徐欣2010: 54)而深得读者喜爱。百度在线机器译本(简称“百度译本”),采用百度在线翻译系统自动生成译文。该语料库实现了英语原文本与三个汉语译本在句子层面的平行对齐,英语原文和中译本则分别进行了词性和分词标注。

2.2 语言特征选定及数据收集

本研究所确定的语言特征均为词汇和句法层面的特征。依据Baker(2000)对译者风格特征的界定和黄伟、刘海涛(2009: 23)及陈芯莹等(2012: 137)提出的用于文本聚类 and 文本作者身份判定的汉语语言计量特征指标,综合语料库翻译学和计量语言学的语言特征概念,确定了词长、句长、标准类符形符比(STTR)及各种实词与虚词比例等16个语言计量特征和风格特征。基于所获数据,考察了三个译本的语言计量特征并分析其成因。

3. 词汇层面特征

由于词汇的频次差异是导致译本语言风格差异的重要原因,加之这些特征易于获取,因此受到计量语言

学和语料库语言学研究的关注(陈芯莹等2012: 137)。本研究借助AntConc3.4.3和Word Smith(5.0)语料库工具,提取了《傲慢与偏见》王译本、孙译本和百度译本词汇层面的语言特征进行考察。

3.1 词汇丰富度

文本中类符(type)/形符(token)比值越高,说明作者的词汇使用越丰富、用词越富于变化(Baker 2000: 250)。表1表明,百度译本的标准类符/形符比(52.02)最高,比王译本(44.01)和孙译本(46.62)分别高了8.01和5.4个百分点,而孙译本则比王译本高出2.61个百分点。数据说明,百度译本的词汇变化和丰富度远高于人工译本,而孙译本的词汇丰富度略高于王译本。这可能是由于百度在线翻译利用海量的网页库资源,挖掘了海量的双语句对的原因。因此,其词库的词汇丰富度和选词范围是任何人工译者所不可企及的,其词汇的丰富度高也是情理之中的事。当然,词汇丰富并不一定意味着译文就好,因为译本质量的优劣并不仅仅由其词汇的丰富度决定。王译本的词汇丰富性不及孙译本,则源于王科一更加忠实于原文的语言结构与用词,而较为忽视在译入语中的艺术风格传达。王认为翻译要极其负责并忠实地传达原作的精神实质(李雪娇2014: 48),而孙则既坚持译文的“忠实准确”,更重视其“生动传神”(孙致礼1991: 61)。从翻译目的论和译者主体性发挥的角度来看,两个人工译本的不同语言特征其实也是译者不同翻译目的和主体性的发挥使然。

将三个译本类符/形符比与王克非、秦洪武(2009: 103)考察汉英对应语料库(Chinese-English Parallel Corpus, CEPC)所得到的汉语翻译文本(47.49)和汉语原创文本(45.19)的类符/形符对比来看,除王译本之外,孙译本(46.62)和百度译本(52.02)的类符/形符比均高于汉语原创文本(45.19)。这也证明Baker和Laviosa等人基于印欧语系的翻译共性中的“简化”特征(1993: 233-250)并非完全适用于所有翻译语言。

表1. 三个译本词汇特征统计数据

译本	王译本	孙译本	百度译本
考察项			
类符	8168	8270	5123
形符	120699	107578	50365
标准类形比(%)	44.01	46.62	52.02
词汇密度	0.8025	0.8086	0.7629
平均词长	1.5049	1.5481	1.5045
形容词比例	0.0498	0.0511	0.0597
名词比例	0.1597	0.1699	0.1598
副词比例	0.1276	0.1244	0.0857
代词比例	0.1503	0.1416	0.1750
助词比例	0.0759	0.0729	0.1074
连词比例	0.0326	0.0324	0.0425
惯用语比例	0.0127	0.0154	0.0042

3.2 词汇密度与词长

词汇密度与平均词长是译文语言风格和可读性的重要指标。从这两个数据来看,三个译本之间差异较

小。百度的词汇密度略小于王、孙译本,说明两个人工译本信息密度相当,亦即阅读难度相当。而百度译本则由于助词和连词等虚词比例较高而消减(offset)了实词的频次,从而使其词汇密度略低于人工译本。三个译本的词长相差无几,说明各个译本的词汇难度相当。而肖忠华(2012:63-69)考察 LCMC 和 ZCTC 语料库得到的汉语原创语言和汉语翻译语言的词汇密度分别为 66.93% 和 61.59%,词长分别为 1.57 和 1.59(见表 1)。数据表明,王、孙和百度译本的词汇密度均高于汉语原创语言,而平均词长基本相当。因此,就词汇而言,汉语译文并没有呈现 Baker 和 Laviosa 等人所描述的“简化”特征(肖忠华 2012:68)。

3.3 各类词的词频差异

为进一步考察各类词汇在译文中使用的具体差异,我们对形容词、名词、副词、代词、助词、连词和惯用语的频次比例进行了统计,以发挥语料库以小见大的特点,从不同词性的词频差异,充分挖掘和分析不同译本的语言特征,并探究其成因。

就形容词来看,百度译本略高于王译本和孙译本,说明在线翻译对英语原文的形容词具有较好的辨别和对等转换能力。从语料库的观察来看,在线机译无法像人工翻译那样,将原文中的形容词灵活转换为动词或副词等。而词性转换不灵活恰恰是在线机译的薄弱之处,说明其智能程度还有待提高。

一般而言,文学作品中的名词都是对故事情节所发生的社会环境和主题人物所处场所的描写和刻画,所占比例较高。表 1 显示,百度译本中的名词比例与王译本基本一致,而孙译本的则略高。这说明,在线机译对名词的处理和识别转换能力较高,但还是逊色于孙译本对词性的灵活处理。这与孙致礼追求译文的“神似”,在神和形无法兼顾的情况下,“宁舍其形”的翻译原则不无关系(孙致礼 1991:61)。

英语中的副词非常丰富,在句中的位置灵活多变,使用的频次也很高。在汉语中,副词的使用也很活跃,往往用来修饰动词、形容词,也可以修饰全句,可以做状语和补语。副词的使用会使人物或时间的描写更加生动活泼。表 1 显示,百度译本的副词比例只有 0.0857,远低于王、孙译本的比例(0.1276 和 0.1244),说明在线机译对副词的识别和处理能力不够,其译文与人工的相比,远不如后者生动和传神。在这一点上,惯用语是个很好的例证。孙译本(0.0154)惯用语比例最高,王译本为 0.0127 而百度译本只有 0.0042,百度不及孙和王译本的 1/3。但就惯用语翻译来看,在线机译具有一定的识别转换能力,并非都“愚蠢”得一无是处,这一点值得肯定。如下例:

例 1: 原文: She was a woman of mean understanding, little information, and uncertain temper.

王译本: 她是个智力贫乏、不学无术、喜怒无常的女

人。

孙译本: 她是个智力贫乏、孤陋寡闻、喜怒无常的女人。

百度译本: 她是一个意思的理解,女人的信息很少,喜怒无常。

从译例来看,人工译本四字惯用语使用活跃,译文感情色彩相当浓郁。而百度译本尽管有误译且感情平淡,但却能翻译出“喜怒无常”这样的四字惯用语,多少有些出人意料,说明有一定的情感识别和表达功能,却也值得赞许。

表 2. 各语料库前 10 位高频词统计

序号	王译本	频率	%	孙译本	频率	%	百度译本	频率	%	英语原文	频率	%
1	的	5142	4.26	的	4675	3.91	的	8 345	8.57	the	4333	3.53
2	她	3410	2.83	她	2846	2.38	她	3 308	3.40	to	4140	3.37
3	他	2880	2.39	了	2656	2.22	他	2 835	2.91	of	3624	2.95
4	我	2592	2.15	他	2634	2.20	我	2 707	2.78	and	3583	2.92
5	了	2565	2.13	我	2603	2.17	是	2 139	2.20	her	2227	1.81
6	不	2176	1.80	不	2269	1.89	不	1 797	1.85	I	2071	1.69
7	是	1874	1.55	你	1784	1.49	在	1 742	1.79	a	1937	1.58
8	你	1747	1.45	是	1500	1.26	你	1 516	1.56	in	1869	1.52
9	说	1350	1.12	说	1171	0.98	了	1 394	1.43	was	1847	1.50
10	—	1343	1.11	—	1168	0.98	和	1 238	1.27	she	1709	1.39

就代词在翻译汉语中的使用,王克非(2012:114)等认为,与原创汉语相比,汉语译文中代词的“显化”特征非常突出。从王、孙和百度译本中的代词比例(0.1503、0.1416、0.1750)(表 1)来看,在线机译的代词“显化”尤为突出,孙译本略低于王译本。我们知道,从句法形态上来看,英汉语分属于形合和意合语言。汉语原创语言中,词语与句法之间的关系不像英语那样用语法手段来连接,往往可以根据上下文的语境来表达语法意义和逻辑关系,省略不必要的人称指代。而翻译文本由于受到源语形合句法结构的制约,代词往往过度使用而呈“显化”特征。从表 2 列举的原文和译文中前 10 位的高频词来看,百度译本的代词(她、他、我和你)总比例远高于两个人工译本。表 1 和表 2 的数据交叉印证,在线机译对源语上下文语义逻辑的判断和处理能力不足。如下例:

例 2: 原文: **He** had always intended to visit **him**, though to the last always assuring his wife that **he** should not go; and till the evening after the visit was paid, **she** had no knowledge of **it**.

王译本: 班纳特先生尽管在自己太太面前自始至终都说是不想去拜访彬格莱先生,事实上一直都打算去拜访他。等到他去拜访过以后,当天晚上太太才知道实情。

孙译本: 本来,他早就打算去拜见他,可在太太面前却始终咬定不想去。直到拜访后的当天晚上,贝内特太太才知道真情。

百度译本:他一直都打算去拜访他,却一再向他妻子说他不去;直到晚上他去拜访过以后,她不知道它。

从例2可以清楚地看出,人工翻译中,用人称替代了原文中频繁使用的人称代词,使译文更加符合汉语的表达习惯,行文流畅。而百度译文则完全对应直译原文的人称代词,在译文中显化了代词,导致人称代词过度使用,译文不通顺,翻译腔十足。

根据肖忠华(2012:87),汉语译文库ZCTC中的助词包括“的”在词频排序中明显高于汉语母语库LCMC。将表1的助词比例与表2中“的”字比例以及英语原文高频词表结合来看,助词使用确实与肖忠华的发现一样。百度译文以助词比例0.1074和“的”字比例8.57高居榜首,“显化”特征明显。在现代汉语中,“的”字属于结构助词,通常附着在词和词组的后面或前面,用以明示结构关系。但过多使用却会造成句子结构的混乱(黄伯荣、廖序东2011:25)。因此,百度译本中“的”字的高频出现正是其句子结构“破碎凌乱”的原因和证据。

王青、秦洪武(2011:124)发现,连词在翻译文本中的频次比原创文本的高,具有“显化”特征。表1显示,王、孙译本的连词比例相当(0.0326和0.0324),而百度译本的略高一些(0.0425)。但是从表2来看,百度译本中并列连词“和”的使用明显高于两个人工译本。用AntConc3.4.3对三个译本并列连词频次的统计发现,王译本使用并列连词1026次,孙译本仅为876次,而百度译文则高达1901次。并列连词的过度使用是译者受英语形合结构影响所致。而在原创汉语中,作者可以根据上下文的语义逻辑将相关的连词内隐。从一定程度上讲,连词的内隐能力是译文流畅度的重要指标。因此,王、孙两个人工译本对连词的内隐能力远超过百度在线翻译。这表明,在线机译对句子的处理只停留在语言表层特征的转换,而不能很好地判断和呈现内隐在上下文中的意义逻辑,以致译文拗口。如下例:

例3:原文“**But** it is very likely that he may fall in love with one of them, **and therefore** you must visit him as soon as he comes.”

王译本“不过,他倒作兴看中我们的某一个女儿呢。他一搬来,你就得去拜访拜访他。”

孙译本“他兴许会看中她们中的哪一个,因此,他一来你就得去拜访他。”

百度译本“但他很可能会爱上一个人,因此你必须尽快去拜访他。”

从例3来看,王译和孙译均对原文连词but、and和therefore进行了必要地内隐,而百度译文则完全外显化处理,使译文流利度降低。

另外,机译对汉语时态标记语的使用尚需改进。表2显示,在线译本中“了”字的比例低于王、孙译本。然而,现代汉语中的“了”字恰恰是个至关重要的时态标记助词。三个译本中“了”字的频次差异至少说明,在线译

本在对英语源文本的时态识别和标示上,还有待提高。

4. 句法层面特征

译文的平均句长也是译者语言风格和译本语言特征的重要标记(Olohan 2004:78)。因此,我们以句号、问号和叹号为汉语的句子标记,统计了原文本和三个译本的句子数及其平均句长,并考察了标点符号、陈述句、疑问句和感叹句的分布。

不同译者在句法层面的操作,既是其翻译目的和译者主体性的体现,也是其语言风格特征的展示。表3表明,原文共有6043句,百度译文的句子数与原文的完全一致,王译本的句子数也非常贴近原文,而孙译本的句子数则远高于原文本和其他两个译本,其平均句长最短。语料库数据与李雪娇(2014:48)的观察一致,即王译本在句子形式上追求与原文本的对等,导致句子结构容量扩大,平均句长增加。孙在追求译文“忠实准确、生动传神”之外,还要尽量提高语言的“可读性”(孙致礼1991:61)。因此,他在句子层面的操纵较多,标点符号的使用较为活跃,从而使得句子数增加,句子结构容量减小,译文的“可读性”增强。不管是王的扩增翻译,还是孙通过增加句子数来压缩句子结构容量和长度,都是为帮助译入语读者更清楚地理解原文,也是翻译共性中“明晰化”假设的有力证据。然而,百度在线翻译由于对句子语义的甄别和判断能力欠佳,无法有效处理和分析句子结构和意义,拆句析句能力不高,其译文句数自然就等同于原文的句数,译文也就无法体现出翻译语言在语义、语法结构和逻辑关系方面的“明晰化”(肖忠华2012:157)特征。

表3. 三个译本句子特征统计数据

译本	小说原文	王译本	孙译本	百度译本
考察项				
总句数	6043	6377	7266	6043
总字(词)数	122869	196023	184926	161742
平均句长	17.2158	30.9331	25.4510	26.7652
标点比例	0.1879	0.1026	0.1115	0.1172
陈述句比例	0.8441	0.8511	0.8665	0.8406
疑问句比例	0.0735	0.0787	0.0651	0.0766
感叹句比例	0.0824	0.0640	0.0683	0.0824

从表3中陈述、疑问和感叹句的比例来看,百度译文对这三类句子的处理与原文几乎一致,这再次证明其对句子综合处理能力不足,只能逐句翻译。孙译本对原文中疑问句和感叹句的灵活处理稍多一些,也进一步印证了其对于原文句型的依赖相对较少,灵活处理译文句子的能力较强。百度译文的感叹句分布数量与原文一致且高于人工译本,这也说明在线机译在对原文语义情感的识别和处理方面有待加强。

5. 结论

本研究基于自建的由《傲慢与偏见》两个人工译本

与一个在线机译译本及小说原文构成的平行语料库,将计量语言学与计量风格学的方法相结合,从词法和句法层面对比了人机译本中16个语言计量特征及风格特征的异同。结果表明,在线译本在词法、句法、情感表达等方面都远不如人工译本,在代词、助词和连词使用上“显化”明显,在类型比和词汇密度方面“简化”不足,而句法层面无“明晰化”倾向,这与其在词性转换、上下文语境判断、情感识别、时态标记识别等方面的能力不足有关。而两个人工译本在词汇和句法层面也呈现出明显的风格差异,其“显化”、“简化”和句法“明晰化”的程度各有不同,其成因与译者不同的翻译目的和主体性意识不无关系。本研究进一步证实了描写翻译学关于翻译语言共性的一些假设,但也证实某些共性假设不太符合汉语翻译语言的规律,同时还指出了在线机译在词法和句法处理及情感识别方面亟待改进的问题。本研究将基于互联网的在线文学翻译纳入研究范围,在这个“互联网+”的时代,具有较强的现实意义,对翻译研究的视阈有所创新和扩展。但由于语料库规模和语言特征考察范围的限制以及对在线机译的了解有限,本研究仅仅是从语言层面对在线机译研究的一个尝试,还有待后来者开展更为深入的研究。

参考文献

- [1] Austin, J. *Pride and Prejudice* [M]. Oxford and New York: Oxford University Press, 1970.
- [2] Baker, M. Corpus linguistics and translation studies: Implication and applications [A]. In M. Baker, G. Francis E. Tognini-Bonelli (eds.). *Text and Technology: In Honour of John Sinclair* [C]. Amsterdam/Philadelphia: John Benjamins, 1993: 233-250.
- [3] Baker, M. Towards a methodology for investigating the style of a literary translator [J]. *Target*, 2000(2): 241-266.
- [4] Olohan, M. *Introducing Corpora in Translation Studies* [M]. London: Routledge, 2004.
- [5] Tucker, Allen B. Current strategies in machine translation and development [A]. In Sergei Nirenburg (ed.). *Machine Translation* [C]. Cambridge: Cambridge University Press, 1987: 22-41.
- [6] 简·奥斯汀. 傲慢与偏见 [M]. 王科一, 译. 上海: 上海译文出版社, 1980.
- [7] 简·奥斯汀. 傲慢与偏见 [M]. 孙致礼, 译. 上海: 上海译林出版社, 1990.
- [8] 崔磊, 周明. 统计机器翻译领域自适应综述 [J]. 智能计算机与应用, 2014(6): 31-34.
- [9] 陈芯莹, 等. 计量特征在语言风格比较及作家判定中的应用 [J]. 计算机工程与应用, 2012(3): 137-139.
- [10] 冯志伟. 基于语料库的机器翻译系统 [J]. 术语标准化与信息技术, 2010(1): 28-35.
- [11] 胡开宝. 语料库翻译学概论 [M]. 上海: 上海交通大学出版社, 2011.
- [12] 黄伯荣, 廖序东. 现代汉语(增订五版) 下册 [M]. 北京: 高等教育出版社, 2011.
- [13] 黄立波. 基于汉英/英汉平行语料库的翻译共性 [M]. 上海: 复旦大学出版社, 2007.
- [14] 黄伟, 刘海涛. 汉语语体的计量特征在文本聚类中的应用 [J]. 计算机工程与应用, 2009(29): 25-27.
- [15] 蒋跃. 人工译本与机器在线译本的语言计量特征对比 [J]. 外语教学, 2014(5): 98-102.
- [16] 蒋跃, 董贺. 计量特征在人机译文语言风格对比中的应用 [J]. 语言教育, 2015(3): 69-81.
- [17] 李浩, 金文. 机器翻译的前世今生 [N]. 光明日报, 2011-9-28(12).
- [18] 李雪娇. 从译者序分析译本特色 [J]. 河北北方学院学报, 2014(4): 48-51.
- [19] 孙致礼. 我译《傲慢与偏见》 [J]. 解放军外国语学院学报, 1991(4): 58-62.
- [20] 王克非. 语料库翻译学探索 [M]. 上海: 上海交通大学出版社, 2012.
- [21] 王克非, 胡显耀. 基于语料库的翻译汉语词汇特征研究 [J]. 中国翻译, 2008(6): 16-20.
- [22] 王克非, 秦洪武. 英译汉语言特征探讨 [J]. 外语学刊, 2009(1): 102-105.
- [23] 王青, 秦洪武. 基于语料库的《尤利西斯》汉译词汇特征研究 [J]. 外语学刊, 2011(1): 123-127.
- [24] 肖忠华. 英汉翻译中的汉语译文语料库研究 [M]. 上海: 上海交通大学出版社, 2012.
- [25] 徐欣. 基于多译本语料库的译文对比研究 [J]. 外国语, 2010(2): 53-59.
- [26] 张律, 胡东平. 测试环境下基于语料库的翻译显化现象研究——以连接词和人称代词主语为例 [J]. 西安外国语大学学报, 2012(2): 114-117.

基金项目: 本文系陕西省社会科学基金项目“贾平凹英译小说语言特征的语料库考察及对外译介研究”(项目编号: 2015K020) 和教育部2015年社科规划项目“在线机译与人工翻译的语言计量特征对比”(项目编号: 15YJA740016) 的阶段性成果。

作者简介: 韩红建, 西安交通大学外国语学院博士生, 西安交通大学城市学院外语系副教授, 研究方向: 语料库翻译学、计量语言学。

蒋跃, 西安交通大学外国语学院教授, 博士生导师, 研究方向: 语料库翻译学、计量语言学。

收稿日期 2015-12-29
责任编辑 周永平