

翻译技术在远东国际军事法庭庭审记录翻译中的应用^①

◎徐 英

1 引言

随着跨文化交际活动的不断扩大，翻译需求急剧增加，质量和效率要求也不断提高。面对高强度高要求的翻译任务，传统的人工翻译方式早已无法满足翻译产业的需求。于是人们便开始借助技术的力量，寻求对传统翻译作业方式的革命。信息技术、大数据和云计算技术的发展，无疑促进了翻译行业的巨大变化，呈现出技术运用的时代特征（王华树、王少爽，2016）。傅敬民（2015：37）认为“现代意义的翻译，是融合了翻译技术的翻译，是在翻译技术支撑下的翻译。翻译技术已经成为翻译领域不容忽视甚至难以分割的重要组成部分”。为此，笔者结合真实案例，探讨翻译技术在远东国际军事法庭庭审记录翻译中的应用。

2 翻译技术

随着信息技术的发展，计算机新算法以及大数据技术的出现，使机器翻译技术取得了巨大突破，越来越受到人们的关注，也得到了翻译业界的广泛认可。机器翻译技术的应用场景越来越多，已经成为翻译从业人员高效生产的重要途径之一。在人工智能技术盛行的时代，翻译技术的迅猛发展对翻译行业和翻译活动本身均产生了巨大影响。翻译技术的应用改变了传统的翻译模式，提高了译员工作效率，提升了翻译行业的生产效率，掌握翻译技术已成为现代翻译行业从业人员的必备技能。

早在2001年，德国的Frank Austermühl就曾详细阐述了信息化时代对翻译工作在信息技术工具方面的要求，要求译者会主动运用互联网搜索、搜集整合和利用语言与翻译数字资源、术语管理、语料库制作与利用、翻译记忆管理、本地化工具、机器翻译等多项技术。根据这一要求，王华树和王少爽（2016）把翻译技术归结于以下五个方面的

^① 基金项目：浙江越秀外国语学院2020年度校级线上线下混合式教学改革项目“基于多平台资源的线上线下混合在《计算机辅助翻译》课程中的应用”。

技术能力：计算机基本技能、信息检索技能、CAT 工具应用技能、术语能力和译后编辑技能。国际翻译服务流程标准 ISO17100（2015）明确地把翻译技术细分为计算机辅助翻译、机器翻译、术语管理、本地化工具、文字处理软件、编辑校对工具、质量保证工具、桌面排版、语音文字识别软件、写作系统、项目管理软件、内容管理系统、翻译管理系统等技术。因此，作为现代译者，有必要在翻译实践过程中综合运用计算机、信息检索、计算机辅助翻译、术语管理和译后编辑等技术，提高翻译效率，确保翻译质量，以应对日益发展的信息技术和急剧增加的翻译需求对翻译提出的新挑战。

根据这一对现代译者的技术能力要求，笔者结合翻译任务特点，尝试在远东国际军事法庭庭审记录翻译实践中应用相关翻译技术，解决翻译实际问题，提高翻译效率。

3 翻译任务特点

笔者所在翻译团队承接了上海交通大学出版社组织的《远东国际军事法庭庭审记录·全译本》第二辑的英译汉翻译任务，笔者负责其中一册，原文近 1500 页，约 30 万字，13 个文档，包含 13 天的庭审记录。拿到原文电子文档，笔者在仔细阅读和研究后发现原文有四大特点：一是充斥着大量历史、军事和法庭等专有名词和法律术语；二是庭审记录记录的是法官、检察官、辩护律师、被告以及证人在庭上的口头语言，既有法律的严谨表述和严密逻辑，又有口语特征；三是有很多的司法审判程式性表达；四是随处可见从句套从句的长段发言。这些特点也就构成了本次翻译任务的四大难点。

3.1 大量专有名词和法律术语

法庭审理的是日军的侵略罪行，证人证词中会不时提到某个人物、官职、历史事件、地方、机构等，因此产生了大量专有名词，如 President 庭长、Marshal of the Court 法庭执行官、General Staff 参谋本部、Jiro Minami 南次郎、Mr. Comyns Carr 柯明斯·卡尔检察官、the China Incident 卢沟桥事变、Mukden 奉天（今沈阳）等。由于待译文档是以日为单位的庭审记录，一个文档少则大几十页，多则一百多页，且往往对某一位被告或证人的询问要持续数日，同样的这名、官职名、地名等专有名词会反复出现多次。这样，就对这些专有名词的译名统一提出了巨大挑战。另外，作为一种专门领域的职业话语，庭审语言有其特殊的话语单元，如 accused 被告、affidavit 宣誓证词、counsel 律师、direct-examination 直接询问、cross-examination 交叉询问。这些用词，是对法律现象和发展规律的提炼和总结，翻译时要体现庭审语言的专业性、准确性和严肃性（刘祥国，2014），不可随意翻译，更不可不统一。这对于没有法律专业背景的译员来说，无疑也是巨大的挑战。

3.2 法律语言和口头语言互相交织

为了保持庭审的权威性和严肃性，检辩双方律师在询问和辩论过程中所使用的语言

要正式、规范和严谨。因此在表达时，双方律师选用的句子结构要紧凑，表达的意思要完整、精确。而被告和证人在回答律师询问时，不可能使用跟律师相同的语言，很多时候是在毫无准备情况下边回忆边回答，语句往往会简短而又零碎，有时甚至是跳跃式表达。这就使得译员在翻译时必须要在两种完全不同的语言风格之间进行切换，且既要保证译文的可读性，又要保持一定的翻译速度，其难度可想而知。

3.3 司法审判程式性表达多

在法庭上，控辩双方针锋相对，不时提交文件作为证据支持或反驳观点，而法官则要做出相应的裁决，这样就会用到很多司法审判的程式性表达。程式性表达往往结构固定化，而且在庭审记录中出现的频率较高，几乎反复出现于每个庭审日。如“The document is admitted on the usual terms. 根据通常规定，文件予以接受。”在每一篇文档中平均出现 12 次；在记录中，几乎只要是检方或辩方律师一开始发言，便以“May it please the Tribunal,”（尊敬的法庭 / 法庭）开头，使用非常频繁；其他诸如“By a majority the Tribunal upholds the objection and rejects the document.”（根据多数决议，法庭支持异议 / 判异议有效，驳回文件）“The documents are in evidence.”（文件作为证据提交）“The objection is upheld and the document rejected.”（支持异议，驳回文件）“Defense document No. 1306 will receive exhibit No. 2512”（接受辩方文件第 1306 号为法庭证据第 2512 号）“We offer in evidence defense document No. 1005.”（我方提交辩方文件第 1005 号作为证据）等也是常见的程式性表达。这些程式性表达既要体现法律性，即在法庭上说法言法语，要严肃地按照某一特点范式说话，又要体现法庭的程序和规则，什么时候说什么样的话是有规定的（张政，2011）。因此，在翻译这类程式性表达的句子时，不可随意发挥，更不可“翻出花样”。

3.4 长难句多

长难句多也是庭审记录语言的一大特点。在庭审记录中随时可以发现夹带着口语但又不是补充添加信息而导致从句套从句的长段发言，它们既不似书面语那般规范，又不似纯口语般简短，有时候令人非常费解，需要在仔细分析解构句子层次和逻辑关系后，才能够捕捉到真正的信息含义，耗费大量时间。下面仅举两例，以示说明。

例 1. Soon after the above plan was put into practice the China Incident broke out, but this was entirely beyond our anticipation and an uneasiness was felt in securing the necessary funds and materials in fulfilling the plan and we who were concerned in the plan were greatly annoyed and wished for speedy settlement.

在该例中，全句有多达 53 个单词，包含 7 个分句，虽然每个分句也都比较简短，但一口气表达了四个主要信息点，如果不处理好这些信息点之间的逻辑关系，乍一读起来，还是有点难以真正搞清楚说话者所要表达的确切意思。因此有必要逐个整理出每一

个分句的基本含义,然后结合上下文和逻辑层次关系,进行汉语表达。

例2. The merchants whom we interviewed at Osaka were inclined to exaggerate certain abuses of boycott methods as racketeering and blackmailing, and to underestimate or even to deny completely the close relationship between Japan's recent policy towards China, and the use of the boycott as a defensive weapon against that policy.

这个句子也比较长,虽然所用的分句不多,但介词短语比较多且较冗长,信息含量大,导致理解起来比较费力,需要把介词短语内容逐一列出。然后整理出恰当的逻辑关系,才能得出清晰的含义。

总之,这样的长难复杂句子,在远东法庭审判记录中比比皆是,虽然逐词逐句理解起来难度不大,但是要想马上用汉语表达出来,还是要颇费一番功夫。

面对文本如此特色鲜明而又对效率和准确率有着高要求的翻译任务,如果采用传统翻译流程和方法,单靠脑力和手工来应对,译员势必无法按时完成。

4 翻译技术运用

针对以上难点和此次庭审记录内容特点,译者大胆借助有关翻译技术,按照译前处理、建立术语库、机译初稿、运用CAT工具进行译后编辑这样一套现代翻译流程进行翻译。

4.1 译前处理

笔者在仔细查看原文后发现,由于原文档是项目委托方对纸质稿扫描后获得pdf格式文档,再通过OCR文字识别得到的word文档,未经整理和排版,因此在标点、格式排版等方面时常出现问题,如字体是非常用字体、英文前引号后多出空格、冒号后文字断行等。另外,由于英文句点与有些单词缩写后加点同为“.”,计算机辅助翻译软件在断句时会将单词缩写后的点如“Mr.”误认为是句点,影响正常断句。这些问题不仅影响计算机辅助翻译软件在导入文档后的断句处理效果进而影响机器翻译效果,给译后编辑带来困难,而且也会影响译文导出后的版面效果。此外,笔者在对照原始扫描版pdf文档后,发现每一份word文档都因首页缺失部分表格内容而导致文字错乱,非常影响翻译的准确性。于是,笔者根据原始pdf文档补齐表格内容并形成模板,以快速补齐其他文档表格内容。内容不全后,开始着手解决上述格式排版问题。首先,运用Word的高级替换功能,将“^p”(注:“^p”表示段落标记)替换为“:”,一次性解决所有冒号后的断行问题。接着,同样运用Word替换功能,把“ ”替换为“ ”,快速解决英文前引号空格问题;把“Mr.”替换为“Mr”,解决机助翻译软件断句问题。最后,参照委托方提供的《译者手册》,对照其要求对原文档进行了重新排版。

4.2 制作术语表，建立术语库

有效的术语管理可帮助译员规避术语翻译错误的风险，保证术语的专业性、一致性和准确性，确保术语使用规范，保证译文质量（王华树，王少爽，2019）。术语管理能力是翻译职业化、专业化、技术化和现代化的体现（陈元飞，2018）。王华树等认为，术语管理能力应贯穿译员的整个翻译流程，在翻译工作中不可或缺。因此，对于像《远东国际军事法庭庭审记录》这样大型重要历史文献的翻译，在翻译过程中运用翻译技术对术语进行有效的管理和应用，就显得尤为重要。

为此，译者首先根据翻译委托方提供的法律术语表和主要人名汉译表，把出现的专名分为人、地、事件、机构、文献五类名词和部分法律术语，制作成统一的英汉双语对照 Excel 术语表。

为诸如人名、法律术语、法庭程式用语等高频词汇做了统一汉译，如图 1 所示。

English	Chinese	Remarks
Accused	被告	
Defender	辩护人	
Counsel	律师	
Marshal	法庭事务官	
Affidavit	宣誓证词	
Allied Powers	同盟国	
Associate Prosecutor	陪席检察官	
Attorney	(辩方) 律师	
Chamber	法官议事室	
Concurring Opinion	协同意见 (书)	(“书”视语境而加)
Majority Opinion	多数意见 (书)	(“书”视语境而加)
Separate Opinion	个别意见 (书)	(“书”视语境而加)
Conspiracy	共同谋议 (罪)	
Count	诉因	
Conventional War Crimes	普通战争罪	
Crimes against Peace	反和平罪	
Crimes against Humanity	反人道罪	
War Crimes	战争罪	
Cross-examination	交叉询问	

图 1 高频词汇统一汉译表

接下来，利用 SDL MultiTerm 2019 套装工具 SDL MultiTerm 2019 Convert 把上述 Excel 术语表中的术语数据转换成 MultiTerm XML 格式文件，XML (eXtensible Markup Language) 是一种数据存储语言，可使用简单的标记来描述数据，形成可快速识别的数据库。再利用工具 SDL MultiTerm 2019 Desktop 将该 XML 格式文件总的数据库导入创建好的术语库 (MultiTerm Termbase)，形成 sdtb 格式术语库文件 (图 2)。当然，还可以利用一款有 SDL 公司开发的免费小工具 Glossary Converter 直接将 Excel 术语表快速转换成 sdtb 格式文件。这样，术语库就制作完成了。译员可以在计算机辅助翻译软件 SDL Trados 中很方便地对术语进行利用和管理，包括术语自动识别并提示、术语查询和术语修改与添加。

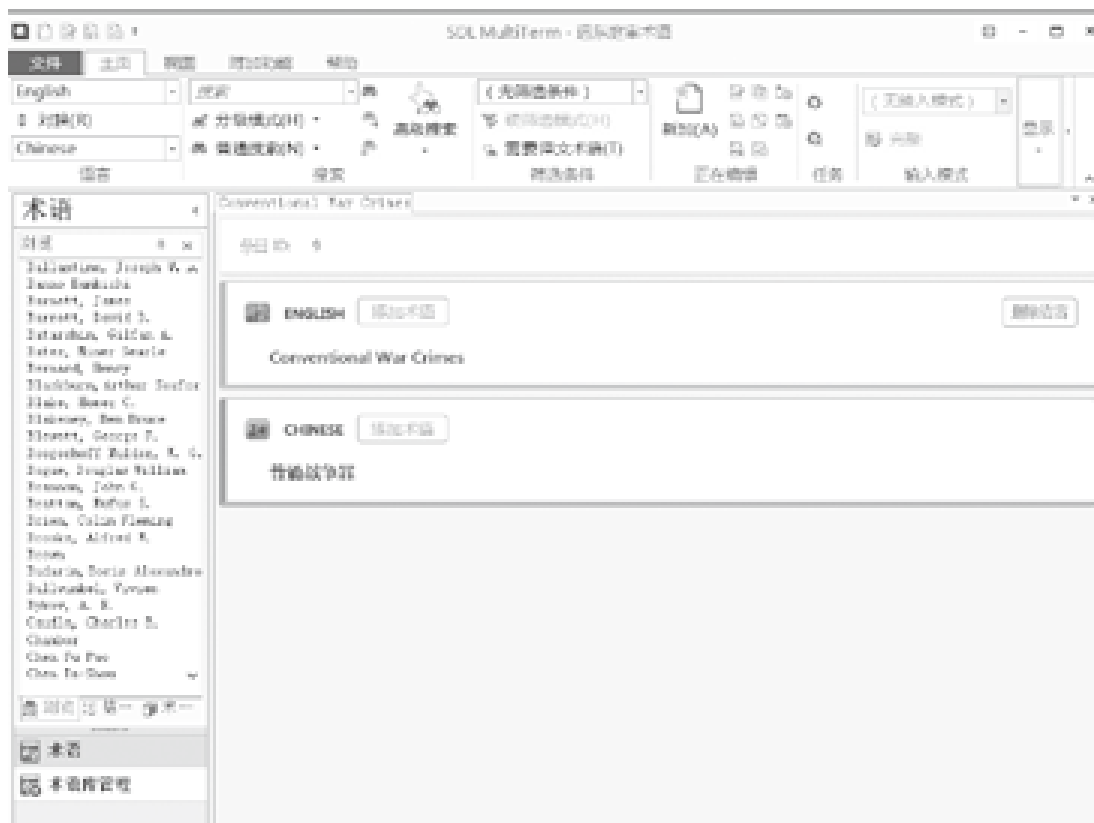


图2 “创建 SDL Trados 术语库”截图

4.3 机译初稿

由于没有前期积累的远东国际军事法庭庭审记录方面的翻译材料，本次翻译任务是历史文献，属首次翻译，因此笔者没有可供利用的现成记忆库。在这种情况下，笔者注意到翻译技术界极力倡导的“机器翻译+译后编辑”翻译模式机器翻译往往适用于特定领域的程式化文本或信息性文本的翻译，如科技文本、法律文本、合同、备忘录、操作手册等。于是，笔者决定采用这种模式，先用在线翻译系统把原文翻译为粗糙译文，再对机器译文进行人工校正和修改，即译后编辑。但是，由于几乎所有在线翻译系统均对免费在线翻译有字符数的限制，无法解决大量文本的免费快速翻译问题。而计算机辅助翻译软件如 Trados 虽然可直接以 API 形式或通过 Tmxmall 的机器翻译插件接入 Google Translator 等机器翻译，但仍然不能免费使用。

经过摸索和研究，笔者最终找到两种方法，解决了无法免费使用机器翻译进行大量文本翻译的问题。第一种方法是借用谷歌或百度翻译的免费网页翻译功能，字数不受限制。具体是将整理好的原文文档保存为 html 网页格式，再用网页翻译功能已设置为开启（图3）的 Chrome 谷歌浏览器（可用谷歌翻译引擎）或 360 浏览器（可用谷歌或百度或有道翻译引擎）打开之前保存为 html 网页格式的原文文档，这时浏览器会弹出“翻译此页？”的对话框，点击“翻译”开始翻译或点击“选项”后设定翻译的语言后进行翻译。若没有弹出“翻译此页？”的对话框，可在该网页上右击鼠标后选择

“翻译成中文（简体）”即可开始翻译。将网页一直往下快速拖动（翻页）直到结束，即可迅速完成整个网页（文档）的翻译。将网页翻译结果复制到 Word 保存，即可完美保留原文档格式排版。



图3 “谷歌浏览器网页翻译功能设置”截图

这种方式虽然可以解决无法用机器翻译引擎免费、快速、批量翻译文档的问题，但只适用于传统的译后编辑方式，无法直接利用计算机辅助翻译软件强大的记忆库和术语库功能来真正提高翻译效率和保证质量。当然，还可以利用另一种翻译技术工具——ABBY Aligner 双语对齐工具将得到的机器译文和原文进行双语对齐，制作成 tmx 通用格式的翻译记忆库文件，再导入 Trados 翻译记忆库中，在 Trados 翻译集成环境下进行译后编辑。

第二种方法是利用一款免费的国产 CAT 工具——Transmate，虽然该软件没有 Trados 那么强大的功能，但其具备了一些基本的计算机辅助翻译功能，如术语库、记忆库管理和机器翻译引擎接入等功能。而且，非常值得一提的是，该软件可以免费接入百度和有道机器翻译引擎。百度机器翻译引擎允许用户每月免费翻译 200 万字符，每个手机号最多可申请 5 个 API。申请好百度翻译 API 后，在 Transmate 中设置机器翻译，创建项目、创建记忆库，然后利用“预翻译”功能，几分钟内就可以完成几十页文档的机器翻译。由于 Transmate 免费版只具备 CAT 的基本功能，在界面编辑、记忆库和术语库利用和质量保证等方面远不及 Trados 强大，故笔者没有选择直接在该软件上进行译后编辑，而是直接确认机器译文使之保存入记忆库后，导出 tmx 格式记忆库文件，再将其作为临时记忆库，运用到 Trados 的远东国际军事法庭庭审记录翻译项目中，结合术语库进行译后编辑。

4.4 利用 CAT 进行译后编辑

有别于直接在 word 文档中对机器译文进行修改的传统“狭义的译后编辑”方式（崔启亮，2014），笔者选择了在由翻译记忆、术语库、质量保证系统和翻译管理系统组成

的 Trados 集成翻译环境中进行译后编辑。把上述经由机器译文双语对齐或在 CAT 工具 transmate 中用百度翻译预翻译而得到的 tmx 格式翻译记忆库文件导入 Trados 中的翻译项目记忆库,再结合术语库,对自动匹配的机器译文进行译后编辑,用修改确认后的定稿译文自动替换原有机翻译文。在翻译过程中,翻译记忆库能够自动对正在翻译的句段进行匹配,根据匹配百分率提供不同的参考译文。译者亦可利用“相关搜索”功能在记忆库中搜索待编译句段中的关键词,查找相似句段的译文,以此保证翻译风格和策略的一致。

作为 SDL Trados 的一大基本功能,术语库为译者在统一和规范专业术语方面发挥了不小的作用。当翻译到相应句段时,术语库会自动跳出相同或高度近似的术语,供译者参考和选用。在遇到陌生且术语库中没有的术语时,译者可充分利用互联网搜索引擎的高级搜索功能和国家图书馆东京审判资源库,解决绝大部分术语问题;还可随时向术语库中添加新术语,供后续使用。这样一来,译者不仅确保了术语的准确性,而且能节省下原本查询或在已有译文中搜索术语所需要的大量时间,专注于文本的理解和翻译,大大提高了效率。

笔者承担的远东国际军事法庭庭审记录翻译项目总共有 13 个独立 word 文档,连续记录了 13 个审判日的庭审过程,经译前整理后页数从 31 到 50 页不等,内容极具关联性,且术语、程式性表达繁多。机器译文不仅能够解决部分术语问题,而且可以把复杂长句按基本语法结构和信息点进行解构,并给出大部分文字翻译,大大减轻了译者查找术语、分析句子解构和打字所花的精力,也节省了时间,大大提高效率如例 3 所示。

例 3. In peacetime there was a very intimate and close relationship between the Supreme Command and the War Minister, as it was through the War Minister that the policy of the government was communicated to the Chief of Staff, and therefore there was no means of conveying the policy of the government to the Chief of the Army General Staff -- to the High Command except through the War Minister or through the Navy Minister, and therefore -- and from that fact there is a very intimate relationship.

百度翻译:在和平时期,最高统帅部和战争部长之间有一种非常密切的关系,因为政府的政策是通过战争部长传达给参谋长的,因此没有办法把政府的政策传达给陆军总参谋长,也就是高层除了通过战争部长或海军部长指挥外,因此——从这个事实来看,有一种非常亲密的关系。

译后编辑:在和平时期,大本营和陆军大臣之间联系紧密,因为政府的政策正是通过陆军大臣才被传达给参谋总长的,除了通过陆军大臣或海军大臣之外,就没有其他途径能将政府的政策传达给陆军参谋总长,再传达给大本营,所以从这一点来看,他们联系紧密。

在上述例子中,虽然百度翻译对于 the Supreme Command 和 the War Minister 等几个专有名词给出了不太准确的译文,但至少给了译者查询确认的方向,而对于 the Chief of Staff 和 the Chief of the Army General Staff 则给出了正确译文。从整体来看,机器译文基

本上呈现了这个复杂长句的大致意思，逻辑也相对准确，只需在用词和语序上稍加修饰和调整，即可得出较为理想的译文。

5 结 语

新时代的翻译工作呈现出越来越鲜明的信息化技术特征，翻译技术的广泛应用势在必行。在远东国际军事法庭庭审记录这样的大型翻译项目中，笔者发挥现代翻译技术优势和译者主观能动性。在运用相关工具对文本进行译前处理的基础上，在以记忆为主要特点的 Trados 集成翻译环境中运用“机器翻译+译后编辑”的工作模式，实行人机互动，既解决了大量法律术语的翻译问题，又省去了译者在传统翻译模式下需要花大量时间去查询和学习法律相关知识的体力和脑力，使笔者能够专注于翻译本身，让原本枯燥乏味的翻译工作变得轻松而富有成就感，从而高效、省力、保质保量地完成翻译任务。但是翻译技术的优势仍然没有充分发挥出来，如在术语库、记忆库的共享和分类标记，在线翻译协作，翻译项目管理，翻译语料库（语言资产）管理和应用等方面还有待进一步探索和尝试。在大数据时代背景下，翻译技术已经成了翻译工作不可或缺的要害，应用翻译技术也是翻译工作者不可逃脱的命运和责任（王华树，2016）。作为翻译工作者，我们应该充分发挥现代翻译技术的优势，主动适应时代发展和市场需求，与时俱进，共同推进现代语言服务行业的发展。

参考文献：

- [1] 王华树,王少爽.信息化时代翻译技术能力的构成与培养研究[J].东方翻译,2016(1): 11-15+73.
- [2] 傅敬民,谢莎.翻译技术的发展与翻译教学[J].外语电化教学,2015(6): 37-41.
- [3] 王华树,李智.人工智能时代笔译员翻译技术应用调查——现状、发现与建议[J].外语电化教学,2019(6): 67-72.
- [4] Austermlühl F.译者的电子工具[M].北京:外语教学与研究出版社,2006.
- [5] ISO17100: 2015. Translation Services—Requirements for Translation Services [EB/OL]. (2019-03-22) [2020-02-22]. <https://www.iso.org/standard/59149.html>.
- [6] 刘祥国.庭审语言特点、分类及策略选择[J].知识经济,2014(13): 28+38.
- [7] 张政.法庭语言特点[N].人民法院报,2011-07-18(2).
- [8] 王华树,王少爽.翻译场景下的术语管理:流程、工具与趋势[J].中国科技术语,2019,21(3): 9-14.
- [9] 陈元飞.职业化背景下译者术语管理之通解——《术语管理概论》评析[J].上海翻译,2018(2): 90-93.
- [10] 王华树,冷冰冰.术语管理概论[M].北京:外文出版社,2017.
- [11] 崔启亮.论机器翻译的译后编辑[J].中国翻译,2014,35(6): 68-73.
- [12] 王华树.大数据时代的翻译技术发展及其启示[J].东方翻译,2016(4): 18-20.

作者简介:

徐英,男,1978年6月生,湖北武汉人,文学硕士,浙江越秀外国语学院英语学院副教授,硕士生,导师,主要从事翻译与翻译教学、翻译技术研究。