

法律术语语料库的创建及应用

李正, 吴越

(曲阜师范大学, 山东 日照 276800)

摘要:法律术语语料库, 是用于法律活动和法律语言学研究的信息库, 其目的是研究法律语言学和收集法律实践案例。目前我国法律数据库建设情况尚不完善, 仍存在诸多问题。文章试图讨论法律术语语料库的建设问题, 分析其背景及应用。

关键词:法律术语; 语料库; 平行文本; 法律实践

中图分类号: H0 **文献标识码:** A **文章编号:** 1009-5039(2018)03-0180-02

1 概述

因法律术语皆来源于西方, 汉译时容易出现术语意义的不对等和缺乏等情况, 从而造成各种问题甚至错误。产生错误的一个主要原因是没有建立完备的法律术语数据库, 而从20世纪90年代起, 语料库和语料库语言学发展迅速, 多种形式的语料库已经在诸多方面发挥了作用。因此, 解决这一问题的一个方法是建立并健全汉英法律语言平行语料库。文章就法律术语语料库的建设问题, 从其背景溯源展开深入讨论, 并对法律数据库的应用提出建设性意见。

2 法律术语语料库的背景

从20世纪90年代起, 随着语料库和语料库语言学的发展, 多种形式的语料库作为经验主义语言研究方法的主要资源, 已经在诸如词典编纂, 语言教学和传统语言研究等方面发挥了作用。

但时至今日有如此强大功能的语料库却主要用于研究和语言本身, 而专门类别和专门语言的语料库寥寥无几。其中的双语语料库, 尤其是涉及汉语的双语语料库的建设, 编码和管理更是少之又少。

法律语言是表达了法律上有拘束力的符号系统^[1]。从语言学的角度看, 法律语言是“领域语言”, 而且可以说是最有现实意义的领域语言, 因而非常值得研究^[2]。众多法律术语因其来源于西方, 在翻译成中文时容易出现术语意义不对等和缺乏。杜金榜(2004)指出, 法律语料库是用于法律活动和法律语言学研究的信息库^[3]。法律语料库是为了法律语言学研究 and 法律实践收集的, 由书面语和口头语构成, 经过计算机处理和存储的, 大量真实的法律文本库集合。法律语料库的建设有着非常明确的目的, 用于描述法律语言和法律实践活动。因而建立完备的法律术语的双语平行语料库, 可以规范法律术语, 收集大量真实法律文本语料和量化数据。不仅能够服务于法律术语翻译实践, 还可以为国内法律英语教学和研究, 特别是法律语言学, 提供帮助。

3 法律术语语料库的建设

3.1 建设步骤

3.1.1 语料采集

挑选一些法律术语类的资料文本等来收集语料。可以通

过网上下载, 使用现存的电子文本, 以及手动输入等方式来收集, 也就是所谓的法律信息库。信息的收集要根据法律术语语料库建设所涉及的以下三个方面: 法律汉语术语语料库, 法律英语术语语料库和法律平行术语语料库。其中, 根据语料库建设的第三条语言从法原则, 即法律语料库中法律汉语语料库和法律英语语料库分别按照法律活动过程划分为立法, 司法, 公安检查和普法四个二级子库^[4]。平行语料库是把两种不同的语言而又相互对应的语句编写进入电脑, 然后对照对比进行研究分析, 找出相互之间所具有的关系, 从而找出同一种内容是怎么用两种语言呈现的。其研究方面包括: 语义, 语法, 语言应用; 相互对比; 两种语言之间的差别; 事实, 实例, 语言环境等。采集时, 要力求覆盖全国所有有效法律, 若不全面, 则语料库就不充分, 不科学, 不可信。

3.1.2 语料规范

从信息库中整理并规范收集到的术语信息, 法律术语必须遵照其真实性原则, 对于由专家定义证实的法律术语可以直接纳入信息库, 而有些模棱两可的术语信息如果不能证实将不能纳入术语库。尤其, 从网络上采集的法律术语语料有时存在乱码、标点符号混乱(全角半角不统一, 英文标点与中文标点混杂)、语料重复甚至错误等问题^[5]。因此更需要行之有效的手段来保证法律术语符合规范, 模糊有歧义的句子在术语库中不能出现。

3.1.3 语料标注

可让机器进行自动分词然后进行对语料库结构, 文献信息, 还有语法的标注。根据COLSEC格式进行语料库的标注; 文献标注要对包括类别, 名称, 作者, 类型, 采集的人, 时间, 地点, 方式等十种具体信息进行标注, 便于后续使用语料时了解来源, 为进一步研究提供方便; 语法标注则是对英语, 汉语, 平行术语语料库进行信息码标注。研究法律语言学, 语料标注为必不可少的重要部分。未经标注的生语料库杂乱无章, 无处溯源, 不能很好地进行进一步深层研究。

3.1.4 检索更新

由于法律术语的特殊性, 检索的基本单位不是字而是词。如今检索软件多种多样, 要选择合适的语料库检索软件, 如果现有的检索软件不能满足需求, 语料库到达高级阶段, 可考虑制造专门的法律术语检索软件。且为了研究法律术语不同时

期的特征,以及在各个年代的发展变化,可设计时间检索模式便于对比。

另外,随着时代的发展,术语库也不可能保持一成不变。应根据时代的变化不断更新术语库,及时淘汰陈旧术语,使术语语料库时刻保持鲜活生命力。

3.2 建设过程中出现的问题

3.2.1 理论研究匮乏

我国鲜有专业人员对法律文本翻译理论的方面进行研究。这导致在进行语料库的相关建设时,由于缺乏相应的理论指导,会产生各种各样的问题。在发表在CSSCI期刊上与法律翻译相关的文章中,鲜有国内研究者针对语料库法律翻译的相关理论做过深入探讨,仅有两篇文章将切斯特曼的对比功能分析方法和沙切维奇的法律翻译交际行为理论作为理论基础,以此来对相应的法律翻译问题进行分析^[6]。从数据上看,国内相关研究者针对法律翻译的理论研究都十分匮乏。对法律语料库的建设产生不良影响。

3.2.2 研究方法内容局限

我国对基于法律语料库的研究方法匮乏。目前我国对法律语料库的研究以定性研究为主,占论文总量的67.6%,定量研究占11.8%^[7],而定性定量相结合的方法则利用的不多。而国外许多相关论文在数据库中只有摘要而无全文,所以其具体研究方不能确定(占文章总量的20.6%)。许多研究者仍然不注意运用混合研究法,而在社会科学研究领域,混合研究法正被大量运用并接受。研究方法的匮乏。会对语料库建设造成不良影响。

3.2.3 利用意识不足

高校相关专业对平行语料库利用意识不足。许多英文法律术语并无定译,通过观察平行语料,能够对语言建立一种较为直观的对比。从而掌握最权威的译法。虽然学生们通过教材中的范例仍然可以获取相应知识,但语料库的优势更为明显,在大量的真实语料中,学生们可以通过自主查询的方式进行自主分析。例如,cross-examination一词,许多词典将其译成“盘问”“盘诘”或“反复讯问”,都译得不对。按照英美法系的审判制度,起诉方和被告方均可要求法院传唤证人出庭作证,在庭上先由要求传证人的一方证人提问,然后再由对方证人提问,也就是起诉方讯问被告方的证人或被告方讯问起诉方的证人,即双方交叉讯问证人,这就是cross-examination,译“交叉讯问”为宜。如果能运用好平行语料库,不但对于学生专业技能的获取大有裨益,反过来也可以帮助我国法律语料库的建设。

4 法律术语语料库的应用

由杜金榜给予法律语料库的定义看来,建设法律语料库的目的就是通过大量真实法律语料揭示法律语言的本质特点,便于进行下一步的语言研究,从而应用于具体的法律实践中。由此可以得出,法律语料库的基本功能是便利法律语言研究及法律实践操作。

法律术语作为一门专门的语体,是法律语言的具体体现^[8]。法律术语的含义较其他而言相对固定,不易造成歧义。此种特殊性决定了其在法律语言中占据重要地位。法律文本自身固有的权威属性决定了法律翻译的严谨性,法律术语的翻译更是不能含糊,有时一个术语的误译会改变整句语义,造成严重后果。然而,法律术语翻译不仅是重点,同时也是难点。不同的法律术语反映不同的法律事务和法律关系。错误使用法律术语,就不能准确反映立法、执法和司法等客观事实和法

制信息,反而会混淆界限,形成误导^[9]。简短精练的法律术语翻译做到功能对等是基本要求,语言对等是更高要求。然而,中外法律的差异性使得两种语言中没有对等词的情况时常出现,法律术语译法不统一,译文混乱,给法律翻译增加了难度,法律术语语料库则解决了这一问题。译者可在法律术语语料库中快速搜索到相应表达,若无具体翻译,也可通过查看相近语料,得出最恰当的译法。

法律术语语料库可用于研究某个法律术语在法律文本中的出现频率,相关搭配及其出现时的语境,根据统计结果确定法律术语的使用方法及适用范围。

例:

造成公民身体伤害或者死亡的其他违法行为

other unlawful acts which cause bodily injury to or decease of a citizen

维护国家安全活动和追查刑事犯罪中的秘密事项

secrets concerning activities for safeguarding state security and the investigation of criminal offences

日常生活中,人们常将“违法”和“犯罪”混为一谈,认为二者并无区别,“违法”即“犯罪”。但通过语料库查找发现,“违法行为”译作“unlawful acts”,“犯罪”译作“criminal offences”,是完全不同的两个概念。

这一功能的最直接应用为编纂法律词典,其大量真实语料为词典词条例句编写提供了来源。同时,语料库的量化研究方式提供的大量可靠的真实数据,便于研究术语在特定时期或体裁中的具体用法。

进行翻译研究时,法律术语语料库为译员和翻译专业学生提供了极佳的平台和参考工具,充分利用语料库可以对术语和习惯表达理解得更到位,做到既省时又精确地翻译,减少错误且效率更高。

法律术语语料库还可以运用到法律语言教材编写及教学中,作为一种强有力的工具和资源,可有效提高教学质量。

5 结论

随着法治社会的发展,人们对法律语言的准确性要求越来越高,建设完备的法律术语语料库刻不容缓。在语料库和语料库语言学发展迅速的大环境下,法律术语数据库的建设会更加顺利,进一步降低工作难度。建设完备后,法律术语库的应用将在很大程度上避免术语不对等或错误的状况,为我国法律事业做出重要贡献。

参考文献:

- [1] 宋北平.对我国“法律语言”概念的反思[J].北京政法职业学院学报,2007(2).
- [2] 宋北平.我国第一个“法律语言语料库”的建设和思考[J].修辞学习,2008(1).
- [3] 杜金榜.法律语言学[M].上海:上海外语教育出版社,2004:95.
- [4] 杨海燕.法律语料库建设设想[J].术语标准化与信息技术,2007.
- [5] 骆嘉宇,徐朱迪,金笑帆,赵若彤.中国海关法律法规中英文平行语料库的构建及计算机辅助术语提取研究[J].科技展望,2016,26(15).
- [6] 张美芳.利用语料库调查译者的文体——贝克研究新法评介[J].解放军外国语学院院报,2002(3):54-57.
- [7] 徐珺,王清然.基于语料库的法律翻译研究现状分析:问题与对策[J].外语学刊,2017(1):73-79.
- [8] 彭楚秋,陶友兰.构建基于语料库的中国“法律术语社区”[J].当代外语研究,2017(1).
- [9] 夏勇.要正确使用法律术语[J].新闻实践,1997(2):35-35.