

# 保密环境下 CAT 技术的应用探索

## ——以党的十九大文件翻译为例

季智璇

(中共中央编译局 中央文献翻译部, 北京 100032)

**摘要:**以十九大翻译工作为例探讨了保密环境下计算机辅助翻译技术的应用模式, 并为其进一步技术改进提供明确方向。计算机辅助翻译软件在无电子稿件且无网络环境的条件下, 可以通过优化翻译记忆库内容及改进翻译项目设置等方式实现一键式检索、协作式检索、参考式检索, 有效提升了翻译效率。通过大会翻译实践也可以看出计算机辅助翻译软件在检索界面优化、译前资料分析、译后术语提取等方面还存在较大改进余地。

**关键词:** 计算机辅助翻译; 保密环境; 十九大; 中央文献翻译

中图分类号: H315.9      文献标识码: A      文章编号: 1008-665X(2018)2-0053-09

### 一、引言

近年来随着信息技术的发展及全球网络的一体化趋势, 计算机辅助翻译(Computer Aided Translation, CAT)的研发取得了长足进步, 相比较尚未成熟的机器翻译(Machine Translation, MT), 其在提高翻译效率、改善译文质量、后期审校编辑等方面具有独特的优势, 并在科技翻译、法律翻译等领域取得了不错的应用成果。

随着我国综合国力的日益强大, 党和国家领导人越来越频繁地出现在世界舞台中央。仅十九大前后短短三个月的时间里, 我国先后举办了金砖国家领导人第九次会晤、美国总统特朗普访华、世界政党高层对话会等多场主场外交活动。面对世界

收稿日期: 2018-02-11; 本刊修订: 2018-02-21

基金项目: 国家社会科学基金重大项目“当代中国重要政治文献多语种数据库建设”(17ZDA108)

作者简介: 季智璇, 翻译, 研究方向: 翻译理论、计算机辅助翻译

越来越渴望听到中国声音、中国智慧、中国方案的新形势，传统的手工作坊式翻译模式显然已远远不能满足中央文献翻译的任务需求。鉴于此，中央编译局中央文献翻译部（以下简称文献部）于 2015 年正式将计算机辅助翻译系统引入中央文献翻译工作，并与传统翻译模式进行了系统整合。

众所周知，计算机辅助翻译提升翻译效率的重要手段之一是网络环境下集成多个机器翻译系统及一键式搜索多个网络语言资源，但中央文献翻译常常需要在保密环境，即无网络环境下开展工作。计算机辅助翻译在无网络环境下能否有效提升翻译效率，本文将以中国共产党第十九次全国代表大会翻译工作为例，分析保密环境下计算机辅助翻译系统的运用模式及其进一步改进措施。

## 二、面向中央文献翻译的计算机辅助翻译体系建设情况

### 1 软件购置及改进情况

文献部结合新时代中央文献翻译工作实际，以网络化管理推动中央文献翻译工作发展，于 2015 年购买了 SDL TradosStudio, SDL MultiTerm 以及 SDL Groupshare 三个组件及计算机辅助翻译系统专用网络服务器，为翻译记忆库管理、术语库管理、项目运行等的网络化运行提供了充分的硬件保障。在实际使用过程中，文献部协同相关技术单位开发了针对不同语言切分语句用的宏程序，提升了中外文语句对齐的效率，开发了翻译记忆库批量转换插件 TMmaker，替代了原有的翻译记忆库转换软件 TMbuilder，提升了翻译记忆库转换效率，并购买术语转换插件 Glossary Converter 提升术语转换效率。

中文	阿文
要破除思维定势，树立与强军目标要求相适应的思维方式和思想观念。	أسلوب وترسيخ الفكري الجود تحظيم يجب كما هدف مع المتطابقين الإيديولوجي والمفهوم التفكير الجيش تقوية ومتطلبات
必须坚持问题导向，坚持战斗力标准，深入研究现代战争特点规律和制胜机理，抓住制约战斗力建设的重难点问题，以重点突破带动整体推进，让一切战斗力要素的活力竞相迸发，让一切军队现代化建设的源泉充分涌流。	لمعالجة المرشد الاتجاه على المثابرة من بد ولا والدراسة، القتالية القدرات معيير وطني، المشكل وأية وقتونها العصرية الحرب لخصائص العبيقة حل على بقوة والعمل، فيها النصر كسب بناء تمنع التي والشائكة الخطيرة المشكلات نقط باختراق الكلي التقدم لدفع، القتالية القدرات القدرات عناصر كافة حيوية تتنجر حتى، هامة لبناء المنافع جميع وتتلفق، ككلمة بصورة القتالية مستفضة بصورة الجيش حصرة
要有针对性地做好思想工作，营造有利于改革的良好氛围，凝聚起改革的正能量，确保部队高度稳定和集中统一，确保改革顺利推进和各项任务圆满完成。	التنهيب عمل إتقان يجب، ذلك جنب وإلى جيد مناخ وتهيئة، ملامن بشكل الإيديولوجي الإيجابية الطاقة وتكتيف، الإصلاح مؤات على الجيش حفظ يضمن مما، للإصلاح ودفع، المركرة والوحدة الدرجة العلي الاستقرار تام بنجاح المختلفة مهامه وإنجاز بسلسلة إصلاحه من خطابه في الجلسة الكاملة لوفد جيش التحرير))
在十二届全国人大二次会议解放军代表团全体会议上的讲话（2014年3月11日），《人民日报》2014年3月12日	الشعبى الصينى المشاركه في الدورة الثانية للمجلس الوطنى الثاني عشر لنواب الشعب الصينى في يوم مارس عام 2014، ونشرته "صحيفة الشعب 11 (اليومية)" في يوم 12 مارس عام 2014

使用 TMbuilder 时，书名、篇章名、作者、会议名称等相关句对属性均需要单独录入。以《江泽民文选（第二卷）》为例，仅录入书名与篇章名就需要 8 小时人工。

图 1 TMbuilder 仅能处理中外文对译内容

中文	阿文	书名	篇章名	原文出处
《全面贯彻落实党的十八大精神要突出抓好六个方面工作》（2012年11月15日），《求是》杂志2013年第1期	وثيقة "التطبيق والتنفيذ الشاملان لروح المؤتمر الوطني الثامن عشر للحزب الشيوعي الصيني" يتطلبان بالضرورة إجابة الأزمات في ستة مجالات الصادرة في يوم 15 نوفمبر عام 2012، التي نشرتها مجلة "البحث عن الحقيقة" العدد الأول عام 2013))	习近平关于全面深化改革论述摘编	一、改革开放是实现中华民族伟大复兴的关键一招	《全面贯彻落实党的十八大精神要突出抓好六个方面工作》（2012年11月15日），《求是》杂志2013年第1期
改革开放是坚持和发展中国特色社会主义的必由之路，所以必须始终把改革创新精神贯彻到治国理政各个环节，不断推进我国社会主义制度自我完善和发展。	الوحيد السبيل هما والانفتاح الإصلاح إن الصينية الخصائص ذات بالاشتراكية للتمسك والإبداع الإصلاح روح جعل فينغي، وتطويرها، الدولة شؤون إدارة حقائق كل دائما تسود نحو بلاندا في الاشتراكي النظام دفع ومواصلة الذائبين والتطور التحسن	习近平关于全面深化改革论述摘编	一、改革开放是实现中华民族伟大复兴的关键一招	《紧紧围绕坚持和发展中国特色社会主义学习宣传贯彻党的十八大精神》（2012年11月17日），《人民日报》2012年11月19日
《紧紧围绕坚持和发展中国特色社会主义学习宣传贯彻党的十八大精神》（2012年11月17日），《人民日报》2012年11月19日	((وثيقة "ضرورة دراسة روح المؤتمر الوطني الثامن عشر للحزب الشيوعي الصيني وتعميمها وتطبيقها بالمحور الوثيق حول التمسك بالاشتراكية ذات الخصائص الصينية وتطويرها" الصادرة في يوم 17 نوفمبر عام 2012، التي نشرتها "صحيفة (الشعب اليومية)" في يوم 19 نوفمبر عام 2012	习近平关于全面深化改革论述摘编	一、改革开放是实现中华民族伟大复兴的关键一招	《紧紧围绕坚持和发展中国特色社会主义学习宣传贯彻党的十八大精神》（2012年11月17日），《人民日报》2012年11月19日
只有社会主义才能救中国，只有改革开放才能发展中国、发展社会主义、发展马克思主义。	على بالإضداد إلا الصين إنقاذ يمكن لا والاشتراكية الصين تطوير يمكن ولا الاشتراكية والانفتاح الإصلاح بتنفيذ إلا والمركسية	习近平关于全面深化改革论述摘编	一、改革开放是实现中华民族伟大复兴的关键一招	《在广东考察工作时的讲话》（2012年12月7日至11日）

使用 TMMaker 时，可直接将不同的书名、篇章名、会议名称等相关属性直填入 EXCLE 表格进行转换。以《江泽民文选（第二卷）》为例，整本书转化时间仅为 2~3 分钟。

图 2 TMMaker 可处理附带多属性中英文对译文档

## 2 人员培训情况

按照全员参与、以新老的思路，文献部以青年译员为重点，先后组织了 2015 年全员基础培训、2016 年“定稿人+CAT 技术骨干”高级培训、2017 年中青年业务骨干高阶培训等专业培训，累计参训人员 130 余人次，部门在职译员至少参加一次，35 次以下青年骨干至少参加 2 次培训。以考促训，组织部门中青年译员参加 SDL 公司组织的 CAT 技能大赛，促进软件使用技能的提升。以讲促训，安排青年技术骨干为大连外国语大学、国际关系学院等院校实习生讲解软件使用技巧，促进青年译员业务能力提升。目前，部门各业务处室均在初译过程中使用计算机辅助翻译软件，部分处室实现了全员参与、全程使用。

## 3 翻译记忆库及术语库建设情况

翻译记忆（Translation Memory, TM）是计算机辅助翻译的核心技术之一（苏明阳，2007）。翻译记忆的工作原理是，译者利用已有的原文和译文建立起一个或多个翻译记忆库，在翻译过程中，系统将自动搜索翻译记忆库中相同或相似的翻译资源（如句子、段落等）给出参考译文，使用户避免无谓的重复劳动，只需专注于新内容的翻译（张宇浩、彭庆华，2014）。

文献部原计划第一阶段建设段对齐翻译记忆库，待技术运用成熟后再开展第二阶段句对齐翻译记忆库建设。但第一期翻译记忆库——江泽民文选段对齐翻译记忆库建成试用后发现每个句对单元字符数过多，导致译文匹配率普遍低于 50%，翻译过程基本与人工翻译的旧模式无差别。利用软件检索词汇时，由于软件检索方式为

模糊搜索, 导致非 100% 匹配的无效结果过多, 毫无参考价值。综合上述理由, 加之翻译记忆库转换插件实现了前文所述的改进, 在第二期翻译记忆库建设中即转为句对齐翻译记忆库, 并取得了良好应用效果。

经过两年半建设, 目前文献部 2000 年之后的著作翻译成果已基本转化为句对齐翻译记忆库资源。截止至十九大召开前, 文献部翻译记忆库资源总计为 43 万句对、2 600 万字 (仅以中文字计), 术语库资源近 7 万条。

表 1 文献部翻译记忆库资源汇总 (截止至十九大召开前)

语种	党代会文件	两会文件	领导人著作	其他文献	外部资料
英文	十八大 十八届四中全会 十八届五中全会 十八届六中全会	1998 年以来 历届两会文件	习近平总书记系 列论述摘编 江泽民文选 邓小平文选 毛泽东选集	中国共产党历史 生态文明 抗战简史 创新驱动发展 十三五规划纲要 求是 (英文版)	习近平谈治国理政
法文	十六大 十七大 十八大 十八届四中全会 十八届五中全会 十八届六中全会	1996 年以来 历届两会文件	习近平总书记系 列论述摘编 江泽民文选 邓小平文选	中国共产党历史	
俄文	十七大 十八大 十八届四中全会 十八届五中全会 十八届六中全会	2010 年以来 历届两会文件	习近平总书记系 列论述摘编 江泽民文选 毛泽东选集	中国共产党历史 党史七十年 “一带一路”建设海 上合作设想	
西文	十六大 十七大 十八大 十八届四中全会 十八届五中全会 十八届六中全会	2007 年以来 历届两会文件	习近平总书记系 列论述摘编 江泽民文选 邓小平文选 毛泽东选集		
日文	十六大 十七大 十八大 十八届四中全会 十八届五中全会 十八届六中全会	2000 年以来 历届两会文件	习近平总书记系 列论述摘编 江泽民文选 邓小平文选 毛泽东选集	中国共产党历史 党史七十年	习近平谈治国理政 中国通信社习近平 总书记相关报道

语种	党代会文件	两会文件	领导人著作	其他文献	外部资料
德文	十六大 十七大 十八大 十八届四中全会 十八届五中全会 十八届六中全会	2002 年以来 历届两会文件	习近平总书记系列论述摘编 邓小平文选（第二卷） 毛泽东选集	“一带一路”建设海上合作设想	习近平谈治国理政
阿文	十七大 十八大 十八届四中全会 十八届五中全会 十八届六中全会	2014 年以来 历届两会文件	习近平总书记系列论述摘编 邓小平文选（第二卷）	习近平在群众路线大会上的讲话 推动共建丝绸之路经济带和 21 世纪海上丝绸之路的愿景与行动 “一带一路”建设海上合作设想	

表 2 文献部术语库资源汇总（截止至十九大召开前）

语种	术语资源
英文	中央文献术语发布平台 汉英政治经济词汇新编 <i>China Daily</i> 汉英词汇摘编
法文	中央文献术语发布平台 汉法政治经济词汇新编 党代会词汇摘编 两会词汇摘编 中译法研讨会（1~26 期）词汇汇总
俄文	中央文献术语发布平台 汉法政治经济词汇新编 时政词汇摘编
西文	中央文献术语发布平台
日文	中央文献术语发布平台 汉日成语摘编
德文	中央文献术语发布平台
阿文	中央文献术语发布平台

### 三、十九大文件翻译工作中计算机辅助翻译软件的配置及应用情况

大会专用电脑处于物理隔绝的单机状态，不仅无法发挥计算机辅助翻译软件调动互联网资源的优势，也无法相互之间构建局域工作网络实现翻译资源的实时共享互通。大会保密组仅提供中文版的纸质稿件，严禁在电脑中录入电子版中文稿件，无法发挥计算机辅助翻译软件利用翻译记忆匹配的核心优势。外文翻译组人员组成复杂，除牵头单位文献部之外，还有很多译员来自外交部、中联部、外文局、新华社、国际广播电台、北京外国语大学、天津外国语大学等相关机构和院校。

考虑到本次大会翻译为第一次在无技术支持（虽然部分涉军机关也在保密环境下使用计算机辅助翻译软件，但在发生故障时可以得到相关技术支持）的保密环境下使用计算机辅助翻译软件，文献部在各语言组自主申请的基础上确定安装 54 套计算机辅助翻译软件，使用人员占到大会外文翻译组成员的 53%。其中，俄文组、日文组、阿文组实现了全员使用，英文组、法文组、西文组、德文组也覆盖了定稿人等关键岗位。

大会工作期间安装了 Trados Studio 和 Multi Term 两个组件，其中 Multi Term 主要安装了术语库检索插件 Widget。大会工作期间既无网络也无电子版稿件，因此文献部对保密环境下计算机辅助翻译软件应用的预期目标为快速检索中外文对译资料。在该预期目标指导下，相关工作人员建立了翻译记忆库检索项目和术语库检索项目。

通过 Trados Studio 建立任意一个翻译项目，利用其检索功能实现翻译记忆库资源的快速检索，检索结果包括中文、外文译文、句对属性等信息。虽然在翻译记忆库界面也就有检索功能，但其检索结果并不具备高亮显示功能，且每次只能检索图 3 中某一个特定翻译记忆库。通过 Multi Term 中的插件 Widget 可快速检索术语资源，检索结果包括中文术语、外文译文、术语出处等信息。

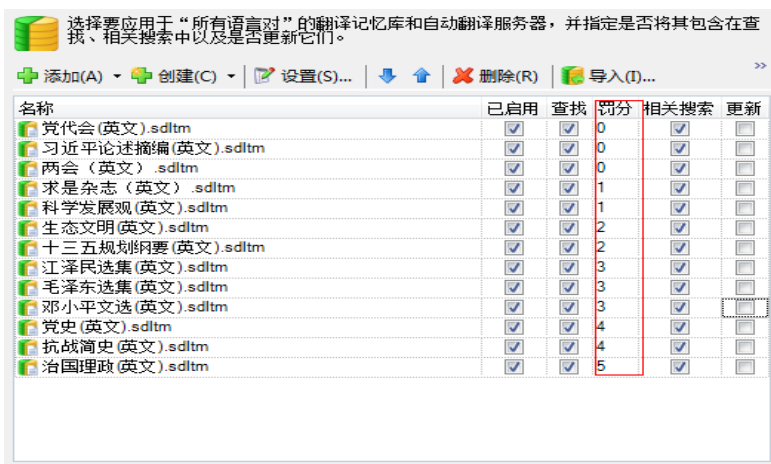


图 3 检索项目翻译记忆库设置情况（以英文为例）

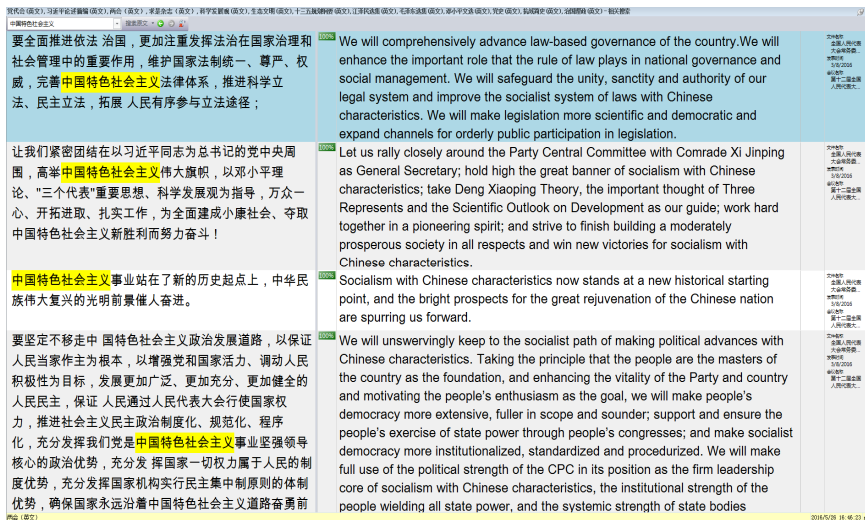


图 4 翻译记忆库检索结果界面（以英文为例）



图 5 术语检索项目术语库设置情况（以英文为例）

传统的 Word 检索方式下，面对少则十几、多则几十的 Word 文档，即使常年从事中央文献翻译的译员也常常会出现漏查的情况，最终选择的译词有很大出入。而临时借调的译员更是无从查起，迷惑之下常常将自己擅长领域的词汇带入中央文献翻译，结果导致后期要花费大量精力去统一译词。通过设置检索项目后，译员仅需要在图 4 左上角检索栏输入检索内容后即可检索图 3 中显示的全部参考资料，大大提升了检索效率。此外，可以通过图 3 显示的界面对翻译记忆库进行评分以实现人工干预，即根据定稿人意见对各参考资料的权重进行统一调整，确保同一语言组内对同一词汇的检索结果相同。图 3 中显示的翻译记忆库设置可根据工作需要灵活调整，在实际工作过程中，由于十九大报告中多处引用《毛泽东选集》

原文，各语言组后期均将《毛泽东选集》翻译记忆库由第二级或第三级参考调整为最优先参考。

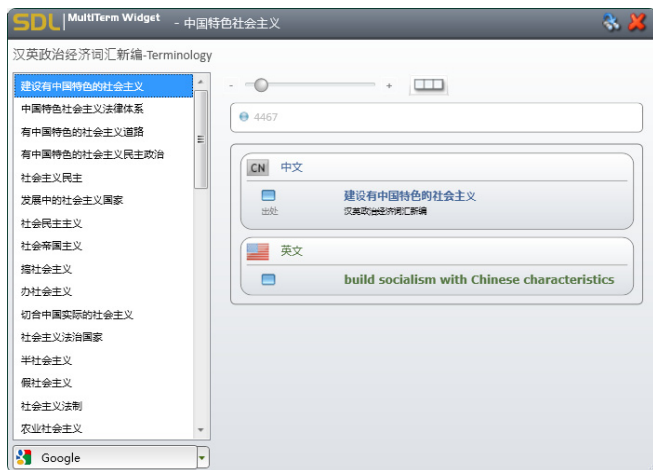


图6 术语库检索结果界面（以英文为例）

大会翻译期间为各语言组均建立了中英对照的检索项目。中央文献中相当数量的政治、经济术语均由英文翻译而来，在将中央文献翻译为其他语言的过程中常常需要先查找其英文原文才能查找到相关语种对应的准确译文。部分具有中国特色的表述，特别是关键表述，如本次大会出现的“习近平新时代中国特色社会主义思想”的英译方案由文献部报送至党和国家领导人得到了认可，因此更具权威性。英文组在近年承担了更多的文献外译任务，其参考资料（13万句对、500万字、14000条术语均为各语言组最多）相比较其他语言组而言更全面。通过为各语言组建立英文检索项目，保障了各语种在关键表述上的一致性，体现了中央政治文献的政治性和权威性。

部分语言组搜集了大量政治新闻类中外文对照资源，并建立了相关翻译记忆库，如日文组搜集的新华社报道及相关中国通信社译文，仅中文字数就达到200多万字。凝练精准的中央文献翻译一直奉行精品战略，虽然译文质量较高，但其每年少则十几万字、多则三四十万字的翻译规模很难保障最新的时政词汇及时得到翻译。政治新闻的展开式报道可以帮助译员准确理解中央文献术语的来龙去脉，且其相关译文解释性的翻译风格也可以帮助译员确定查找更加凝练精准的翻译方案的方向。日文组在相关参考资料库的协助下，最快完成了报告的初译工作。

#### 四、保密环境下计算机辅助翻译存在的问题及改进建议

翻译记忆库与术语库的检索项目需要分别建立，即一键式查询仅分别针对翻译记忆库和术语库。而从实际使用效果来看，很多译员往往仅使用翻译记忆库进行查



询, 很少查询术语库。这其中固然有翻译记忆库查询结果有上下文语境支持, 而术语查询结果中很多缺少例句支持的原因, 但不能实现真正意义上的一键式查询也是一大重要原因。建议计算机辅助翻译软件进一步整合检索界面, 以目前技术成熟程度来看, 建议以术语库查询插件 Widget 整合翻译记忆库查询功能, 实现真正意义上的全资源一键式查询。

大会翻译期间, 部分翻译记忆库在检索过程中被发现存在错误, 即中文与外文不相匹配的现象。由于时间紧、任务重, 加之计算机辅助翻译软件并没有附带错误纠正或错误标记功能, 译员只能选择忽略错误句对。面对每个语言组动辄 10 余万句对的翻译记忆库, 通过全部人工筛查的方式去纠正错误不仅效率低下, 且容易发生疏漏。建议计算机辅助翻译系统开发翻译记忆库错误标记功能, 方便译员对利用过程中发现的错误及时标记, 并由技术维护人员在后期根据标签进行更正, 确保翻译记忆库内容的准确性和权威性。

对于大型翻译项目而言, 译前准备是一项非常重要的工作。面对中央文献翻译精品战略导致大会期间参考资料不足的困境, 如大量搜集政治新闻形成参考资料库固然是一种可行办法, 但依然不能做到对大会文件精准施策。考虑到大会文件的特点, 通过搜集党和国家领导人近五年间的讲话并进行高频词汇分析, 针对出现频率最高的 50 至 100 个术语在译前准备中制定出翻译方案。这对于提升大会翻译效率而言将是更加治本之策。建议计算机辅助翻译系统开发文件预分析功能, 针对特定文件或文件群开展大数据分析, 筛选高频语句或术语。相关成果将不仅对翻译策略的制定具有重要参考价值, 更将推动语言学、政治学等相关学科的进一步发展。

翻译的最终目的在于对外传播。相比较长篇累牍的文章, 短小精悍的术语更适合当下以新媒体为主流传播渠道的媒体环境。过去术语提取工作主要依靠资深译员人工选取, 不仅耗费时间长, 而且术语的权威性常常遭到相关领域专家的质疑, 因此, 如何利用大数据快速提取大会外文术语是中央文献翻译面临的一个紧要课题。相关合作软件公司提供的术语提取插件并不完善, 在之前的翻译项目中并没有得到任何有效结果, 因此, 本次大会未尝试使用计算机辅助翻译软件提取术语。建议计算机辅助翻译系统增强术语提取功能, 在快速准确的前提下提供具有详实数据支撑的术语表, 提升大会翻译的海外传播效果。相关成果将不仅对今后中央文献翻译实践产生帮助, 也可以结合语料库语言学等对翻译理论和翻译批评的发展产生推动作用。

#### 参考文献:

- [1] 苏明阳. 2007. 翻译记忆系统的现状及其启示[J]. 外语研究, (5): 70-74.
- [2] 张宇浩, 彭庆华. 2014. 浅析计算机辅助翻译中的译者主体性[J]. 长春工业大学学报(高教研究版), (1): 142-144.

### **Answering-questions Work in the Translation of the 19th CPC National Congress Documents: Research on the Political Language and Working Mechanism**

XIONG Dao-hong

**Abstract:** The work of answering questions, an important part of the translation work of general political meetings in China, provides reference to the questions relevant to the understanding of Chinese texts during the process of translation. These questions are considered valuable empirical materials for academic research. An analysis of the subjective and objective reasons of the questions put forward during the translation of the 19th CPC National Congress documents would help to understand the key points and difficulties in political documents translation, may also reveal certain language features of contemporary Chinese political documents in political linguistic sense, and contributes to improving the working mechanism of translation. The translation of political documents consists of two parts, one is from the political language to the ordinary language, the other is the ordinary language to the target language. In this process, the translator always undertakes the mission of great political significance.

**Key words:** the 19th National Congress of the Communist Party of China; political documents translation; answering-questions work; political language; working mechanism.

### **The Application of Computer Aided Translation Technology in a Translation Environment of Strict Confidentiality**

JI Zhi-xuan

**Abstract:** This paper discusses the patterns and improvement of computer aided translation (CAT) in confidentiality environment based on the translation of the 19th CPC National Congress documents. Without electronic manuscripts and access to the Internet, CAT software can realize one-click search, coordinated search and reference search to improve translation efficiency by optimizing the Translation Memory and improving the setting of translation project. Through the translation application, the paper suggests that CAT software should improve in search interface, data and term extraction.

**Key words:** CAT; environment of confidentiality; the 19th National Congress of the Communist Party of China; CPC documents translation

### **Meaning Negotiation and Interaction in Discursive Practice: A Case Study of a Press Conference of Chinese and American Leaders**

LIU Li-hua & SUN Ju

**Abstract:** This study, based on the discourse studies, explores the discourse models of Chinese and American leaders' interaction in a press conference. It is found that the President of the United States, as the leader of the super power in the world, manipulates the flow of the meaning in the interaction. He deliberately chooses the topics that are against China but in favor of America to consolidate China's promise, obligation and responsibility to America and the world. He also employs the exclusive pronoun *us* and