

# 计算机辅助翻译

## 理论与实践

张霄军 王华树 吴徽徽 编著

陕西师范大学出版社有限公司  
Shaanxi Normal University General Publishing House Co., Ltd.

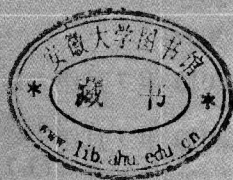
# 计算机辅助翻译

## 理论与实践

张霄军 王华树 吴徽徽 编著

本书获陕西师范大学校级质量工程教材建设项目资助

陕西师范大学出版总社有限公司  
Shanxi Normal University General Publishing House Co., Ltd.



图书代号: JC13N0760

### 图书在版编目 (CIP) 数据

计算机辅助翻译: 理论与实践 / 张霄军, 王华树, 吴徽徽编著. —西安: 陕西师范大学出版总社有限公司, 2013.7

ISBN 978-7-5613-7151-0

I. ①计... II. ①张...②王...③吴... III. ①自动翻译系统—研究 IV. ①TP391.2

中国版本图书馆 CIP 数据核字 (2013) 第 143969 号

### 计算机辅助翻译: 理论与实践

张霄军 王华树 吴徽徽 编著

---

责任编辑 曾学民  
责任校对 涂亚红  
封面设计 王鑫瑞  
出版发行 陕西师范大学出版总社有限公司  
社 址 西安市长安南路 199 号 (邮编 710062)  
网 址 <http://www.snupg.com>  
印 制 兴平市博闻印务有限公司  
开 本 787mm×1092mm 1/16  
印 张 20.5  
字 数 200 千  
版 次 2013 年 7 月第 1 版  
印 次 2013 年 7 月第 1 次印刷  
书 号 ISBN 978-7-5613-7151-0  
定 价 49.20 元

---

读者购书、书店添货或发现印装质量问题, 请与本公司营销部联系、调换。

电话: (029) 85307826 传真: (029) 85303622

# 目

# 录

第一章 翻译技术概论	1
本章导读	1
1.1 翻译与技术：译员工作模式的改变	1
1.1.1 译员工作平台的嬗变	2
1.1.2 网络翻译平台	3
1.2 “人助机译”与“机助人译”	6
1.2.1 人和机器在翻译活动中的分工	7
1.2.2 “机助人译”与“人助机译”的哲学透视	10
1.2.3 人助机译和机助人译的融合	13
1.3 本书内容介绍	14
思考题：	16
第二章 文本处理系统	17
本章导读	17
2.1 文本获取	17
2.1.1 字符与编码	17
2.1.2 扫描输入	21
2.1.3 语音输入	23
2.1.4 文件格式	26
2.1.5 翻译中的文件类型	28
2.1.6 排版文档	32
2.1.7 文件格式转换	33
2.2 文本分析	33
2.2.1 词频统计与字/词数计算	34

2.2.2 索引与搭配	39
2.2.3 文本标注与标记规范	49
思考题:	53
第三章 术语管理系统	54
本章导读	54
3.1 术语的识别与抽取	55
3.1.1 术语识别	55
3.1.2 术语抽取	59
3.2 术语的存储与检索	63
3.2.1 术语存储	64
3.2.2 术语检索	69
3.3 术语管理与术语管理工具	72
思考题:	74
第四章 翻译记忆系统	75
本章导读	75
4.1 原理: 文本匹配检索	75
4.1.1 切分	76
4.1.2 匹配	77
4.1.3 匹配算法	81
4.2 创建: 文本对齐处理	88
4.2.1 平行语料预处理	88
4.2.2 自动对齐方法研究	91
4.2.3 平行语料加工	94
4.3 翻译记忆系统的优缺点	103
4.3.1 翻译记忆系统的主要优势	103
4.3.2 翻译记忆系统的局限性	103
思考题:	104
第五章 计算机辅助翻译新技术	105
本章导读	105
5.1 本地化技术	105

5.1.1	本地化、全球化和国际化	105
5.1.2	翻译与本地化	106
5.1.3	本地化工具	108
5.1.4	本地化过程	109
5.1.5	可视化本地化	113
5.1.6	本地化技术示例	119
5.2	云计算技术	125
5.2.1	云计算	125
5.2.2	云翻译	128
5.2.3	SaaS 翻译技术	130
5.2.4	云翻译技术示例: SDL BeGlobal	132
	思考题:	134
第六章	术语管理实践	135
	本章导读	135
6.1	个人译者/翻译团队的术语管理	136
6.2	计算机辅助翻译工具的术语管理模块	145
6.2.1	D é j à vu 术语管理模块	145
6.2.2	Wordfast 术语管理模块	163
6.3	企业级术语管理	169
6.3.1	大型企业术语管理存在的问题	169
6.3.2	企业级术语管理解决方案	170
6.3.3	企业级术语管理系统	171
6.4	基于网络的术语管理	202
6.4.1	常见的基于网络的术语管理系统	202
6.4.2	免费和开源的术语管理系统	212
第七章	SDL Trados Studio 2009 实践	215
	本章导读	215
7.1	SDL Trados 2009 主要功能介绍	215
7.2	基本操作	218
7.2.1	项目情况介绍	218

7.2.2	创建翻译项目	219
7.2.3	项目批任务处理	223
7.2.4	查看分析报告	224
7.2.5	翻译及审校操作	227
7.2.6	翻译审校工作	232
7.2.7	定稿	233
7.3	翻译记忆库创建及管理	234
7.3.1	翻译记忆库创建	235
7.3.2	在翻译过程中更新字段	237
7.3.3	翻译记忆库的维护和管理	238
7.4	对齐已翻译文档	244
7.4.1	新建对齐已翻译文档项目	244
7.4.2	匹配文件对	248
7.4.3	导出匹配结果	250
第八章	Wordfast 实践	251
内容导读		251
8.1	Wordfast 基本介绍	251
8.2	翻译案例分析	258
8.3	Wordfast 基本操作	258
8.3.1	创建翻译记忆库	258
8.3.2	创建翻译术语库	260
8.3.3	创建翻译项目	263
8.3.4	预翻译 (MT)	265
8.3.5	翻译编辑	267
8.3.6	生成译文	272
第九章	其它计算机辅助翻译系统	274
本章导读		274
9.1	B/S 架构的 CAT 系统	274
9.1.1	Lingotek	275
9.1.2	Google Translator Toolkit	280

9.1.3 Wordfast Anywhere .....	286
9.1.4 小结 .....	292
9.2 开源计算机辅助翻译工具 .....	292
9.2.1 Okapi Framework .....	296
9.2.2 OmegaT .....	300
9.2.3 Transolution .....	304
9.2.4 Virtaal .....	306
9.2.5 小结 .....	308
附录一: .....	309
ASCII 字符集 .....	309
附录二: .....	310
建立术语数据库有关的国家标准和国际标准 .....	310
附录三: .....	314
Wordfast 常用快捷键 .....	314
参考文献 .....	316





# 第一章 翻译技术概论

## 本章导读

传统的翻译服务是基于译员的个人知识的，因此其翻译平台一般是纸笔媒介或者个人电脑；而现代翻译服务则是基于团队远程协作的，因此必须依靠一个网络化翻译服务平台。在现代翻译服务中，知识的管理与共享是依靠网络平台实现的。在这个统一的平台上，管理者、业务人员、专业译员和专业译审等成员之间是通过网络，应用统一的翻译管理系统和借助特定的计算机辅助软件来共同完成翻译项目。

在翻译技术领域，“机助人译”“人助机译”“机器翻译”和“计算机辅助翻译”等几个术语的使用非常混乱，术语的不确定会造成读者对术语表述对象理解的偏差。本章从术语的解读入手，分析了“机助人译”和“人助机译”在翻译活动中的功能区分，又从人工智能哲学和技术哲学的角度对这两种翻译技术做了哲学分析，指出在新的翻译环境下人助机译和机助人译的融合趋势。最后对本书各章主要内容做了介绍。

### 1.1 翻译与技术：译员工作模式的改变

目前，翻译服务越来越产业化，翻译已经成为了一个产业。在整个翻译服务流程中，翻译服务供应商和翻译客户之间的交流和信息传递是必不可少的，而在翻译服务提供商内部工作流程中，翻译项目负责人、译员与译审之间的交流和信息传递也是必不可少的，同时，翻译项目负责人还要和市场跟单员以及多个译员、多个译审进行协调和交流。上述各种翻译服务参与者之间的沟通交流、信息传递和协调合作都必须在一个统一的网络化的工作平台上进行，否则，翻译项目的管理就会产生混乱，从而造成翻译成本增加，翻译质量下降，整个翻译服务流程就会费时费力。

“平台”原意是指“生产和施工过程中为操作方便而设置的工作台”(《现代汉语词典》), 英语中把“平台”称为“bench”, 那么“工作平台”就为“workbench”。和大部分脑力工作者和办公室工作人员一样, 翻译工作者(即译员)的工作平台也经历了纸笔平台、PC 平台和网络平台三个阶段。在整个翻译服务活动中, 译员的工作居于核心位置(朱宪超, 韩子满, 2006:27-28), 因此, 译员工作平台的嬗变也就是整个翻译活动工作平台的发展。

### 1.1.1 译员工作平台的嬗变

#### (1) 纸笔平台阶段

纸笔平台阶段, 译员的翻译服务是通过纸和笔来完成的。译员接到的是手写的纸质的原文文本, 交出的是手写的纸质的译文文本, 整个翻译服务流程中所用到的翻译介质只有纸和笔。这是传统翻译服务的主要工作形式, 这种纸笔工作平台持续了上千年, 直到有印刷体的书本出现。这时翻译服务的工作流程就变成了: 译员接到的是印刷的纸质的原文文本, 交出的是印刷的纸质的译文文本或者手写的纸质的译文文本(之后由翻译客户自行印刷)。但这种工作模式也没有脱离纸笔平台的窠臼, 因此, 也属于纸笔平台阶段。

#### (2) PC 平台阶段

真正对纸笔平台模式产生冲击的是电子计算机的出现, 译员的翻译工作进入了 PC 平台阶段。上世纪 90 年代初, 个人电脑的普及使得我国的翻译工作者经历了一次集体“换笔”行为, 即几乎所有的译员都使用 PC 作为翻译工作的平台。这个阶段译员的基本工作模式是: 译员接到的是打印的纸质的原文文本, 交出的是打印的纸质的译文文本。其实, 这种模式还是传统翻译服务模式的延续。后来, 随着 Internet 的出现, 纸质文本是否需要就成了译员和客户的个体取舍行为, 通过电子邮件原文和译文多以电子文本的形式进行交流。但这还不是真正意义上的“网络平台”, 因为在此阶段, PC 在整个翻译流程中依然是主体平台, 互联网只是文本传输的一种途径。

#### (3) 网络平台阶段

在纸笔平台阶段和 PC 平台阶段, “翻译服务供应商”这个概念比较模糊, 大部分的翻译活动似乎也不需要这个角色。然而, 随着语言需求的增加和国际交流的加剧, 翻译服务的需求变得越来越迫切。根据国际权威机构对世界翻译市场的

调查,全世界翻译市场的规模在1999年只是104亿美元,而在2005年则达到了227亿美元。目前,我国翻译能力严重不足,尽管翻译市场的规模已经超过了100亿人民币,但是现有的国内翻译公司只能消化10%左右(Austermuhl,2006:ix)。由于无法消化大量从国际上传来的信息流,就有可能在国际竞争中失去大量机会。在这种情况下,传统的手工翻译和单枪匹马式的个人翻译已经难以满足不断增长的翻译需求,“翻译服务供应商”应运而生,译员的翻译活动发生了质的改变——翻译手段辅以电子工具,翻译方式改为团队协作。译员赖以辅助的电子工具和团队队员之间的协调都需要翻译服务供应商提供,此外,和翻译客户的沟通交流以及对译员翻译质量的监控都需要翻译服务提供商来完成。译员虽然在翻译活动中依然占据核心地位,但起主导作用的却已经不是译员,而是翻译服务供应商。翻译服务供应商要很好地扮演这个角色,就需要一个统一的管理、翻译和监督的工作平台来做道具,而且这个平台是建立在网络协同作业基础上的,译员的翻译服务进入了“网络平台”时代。

### 1.1.2 网络翻译平台

一个满足现代翻译需要的网络化服务平台应该满足以下要求:(1)确立翻译项目诸参与者的地位;(2)支持译员所使用的辅助翻译工具;(3)便于知识积累和资源储备;(4)实现即时的协同作业;(5)加强翻译流程控制。拥有一个这样的网络化服务平台的翻译服务供应商还应该通过ISSO9001认证体系,因为,从客户视角来看,最理想的“翻译提供者”应该是具备ISO9001:2000认证的翻译公司和翻译中介。(Samuelsson-Brown,2006:34-41)

#### (1) 确立翻译项目诸项目参与者的地位

现代翻译服务都是以“翻译项目”的形式进行的,一个完整的翻译项目的参与者至少应该包括客户、项目负责人(一位)、跟单员、译员(一位或多位)、译审(一位或多位)以及技术支持。在现代翻译网络平台上,译员依然居于翻译活动的核心位置,因为不论采用什么样的翻译平台,也不论利用什么样的翻译工具,译员都是原文的接受者和译文的产出者。译员的核心地位不会发生改变,但译员在整个翻译项目中不再居于主导地位,因为他必须接受项目组的技术支持为他提供的技术支持和项目负责人指定的审校对他翻译质量的监督,他也不再直接和客户联系,中间由负责人安排的跟单员负责联系。由此可见,居于翻译活动主导

地位的应该是项目负责人，项目负责人还负责译员的考核和翻译进度的监督，另外，对于审校和技术支持的工作他也能通过网络平台进行实时监控，对整个翻译流程进行控制，必要的时候，他还要负责和客户进行沟通。以上翻译项目参与者之间的关系如图 1.1 所示：

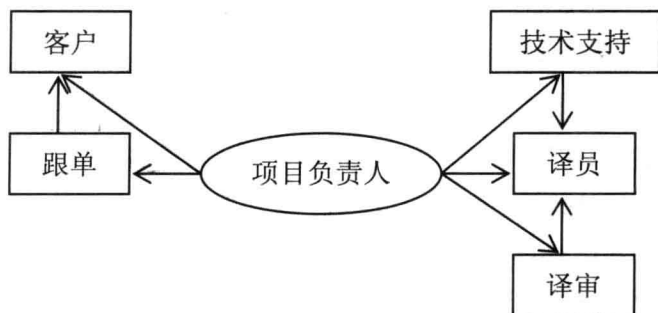


图 1.1 现代翻译项目诸参与者之间的关系

### (2) 支持译员所使用的辅助翻译工具

“全自动高质量机器翻译（fully automatic high-quality translation,FAHQT）”的愿望和目标因为困难过大，已经慢慢被放弃了。越来越多的译员开始使用各种翻译工具进行翻译工作，以满足现代翻译的需要。这些翻译工具统称为“计算机辅助翻译（computer-aided translation,CAT）”，可以分为“人助机译（human-assisted machine translation,HAMT）”和“机助人译（machine-assisted human translation,MAHT）”两种类型（陈群秀，2007:105-113）。在 HAMT 中，一般由译员来进行译前编辑或译后编辑，或者进行人机交互，主要的翻译工作基本上是由机器来完成的；在 MAHT 中，译员借助于电子词典和翻译记忆等电子翻译工具，源语言的分析 and 翻译工作都是译员手工完成的。不论是 HAMT 还是 MAHT，在一个翻译项目中，文本的编辑和工具的使用都要能在一个统一的平台上使用，这样才有利于译员之间的合作和译员与审校的配合。

### (3) 便于知识积累和资源储备

“翻译记忆（Translation Memories,TMs）”软件能够保存和重复使用译员已经翻译好的译文。如果有新的资料需要翻译，可以使用原来存储在翻译记忆里的译文。原来的译文与新的资料之间要进行匹配，翻译记忆软件可以根据匹配程度的不同来决定翻译策略。翻译记忆是翻译服务供应商重要的知识资产。应用 TM 技

术,后台记忆每一个新的翻译结果,这样使用户能最大限度地减少重复工作。作为知识资产的翻译记忆库,保存的译文越多,降低的成本也就越多。TRADOS公司的翻译记忆系列产品 Translator's Workbench 就是 TMs 软件中的典型。平行语料库(parallel corpus)也是语言资源储备的一项重要内容,除了收集以前翻译好的译文之外,可以被计算机程序检索的、具有一定规模的对应语料的收集也是必需的,它可以帮助译员判断译文是否正确(Pearson,2007:15-24)。

术语管理是计算机辅助翻译的一大优势。在专业技术文献的翻译过程中,最令人头疼的问题是术语的翻译。为了找到准确的术语,译员往往得花大量的时间。翻译网络平台中应用计算机辅助术语管理系统可以大大缩短译员查询术语的时间。

#### (4) 实现即时的协同作业

在一个翻译项目中,项目参加者之间的交流是即时的,翻译项目的进行和完成是协同作业的。

在传统的翻译活动中,同一项目的不同译员之间的翻译是“串行”的、孤立的。不同译员的翻译最后由统稿者统一定稿,译员之间缺乏交流和沟通。翻译项目虽然是由团队“合作”完成的,但不是团队“协作”完成的。在网络翻译平台上,译员可以通过短消息的方式实现“并行”的、即时的交流和联系,互通有无,完成翻译的协同作业。

传统的翻译活动中,译员和审校的工作是不同步的,一般是译员翻译完,译审才能审校。有时是译员在翻译完一部分时就要交稿让译审来审读,当他翻译其他部分时前面翻译的就无法参考了。译员在翻译过程中即使有翻译难点也不能得到译审的即时帮助。在网络翻译平台上,译员和译审能即时通讯,随时交流翻译问题。

质量是翻译的生命,质检在翻译活动中起着保障生命线的作用。翻译的质检一般采用“取样法”和“反馈法”,前者是从译员业已完成的译稿中随机抽取一定量的译文进行审查检验,后者是从客户对翻译产品的反馈意见中评判译员的翻译质量。但这两种传统的翻译质检方法都是在翻译项目中完成,起码是译员交稿之后进行的,不能有效地检验出翻译过程中的译文质量缺陷,不能即时地监控翻译质量。在网络化翻译平台上,项目负责人可以随时采用取样的方法对正在进行的翻

译项目进行质量监控，也可以随机提取译文样本让客户进行意见反馈，根据反馈意见在线指导译员的翻译。

### (5) 加强翻译流程控制

对翻译流程的控制包括译员的选择、翻译进度的监控、翻译质量的监督、翻译质量的评价等。

现代翻译中，每一个翻译项目需要的译员可能不止一位，任务紧时甚至需要十几位译员来共同完成一个项目。译员的工作地点也不固定，通过网络在家里就可以完成翻译。这就对翻译项目负责人挑选译员提出了要求，一般来说，一个项目中得有一个专职的译员，其余译员可以通过网络平台从译员库中进行选择。在试译阶段，负责人根据平台上反映出来的各位译员的翻译情况进行正确的译员选择。项目进展过程中，负责人可以根据每位译员的翻译进度适当地增减译员数量。遇到翻译质量有起伏的译员，负责人可以通过平台对齐进行即时的指导，如果不能保证质量，就用备选译员进行替换，保证项目的正常进行。对翻译质量的评价，翻译平台有一套自己的评价体系，可以从量化的角度科学客观地对译员的工作进行肯定和批评。

应用网络翻译平台进行现代翻译服务具有以下优点：(1) 能够实现知识积累和管理的网络化，使翻译参与者共享网络知识资源，从而完成跨地域协同作业；(2) 能够实现翻译产业的工业化，大大提高翻译服务的质量和效率，加强对翻译流程的控制，将传统的“翻译作坊”转变为现代“翻译工厂”；(3) 能够实现翻译过程的即时化，做到“译审同步”，从而改造传统的翻译生产过程，有效解决了传统翻译过程中术语不统一、表达不一致等难题，使译员与审校者之间能够即时交流。

现代翻译服务公司的核心竞争力来自信息技术和管理，而网络信息技术则是信息技术的重中之重，管理包括翻译知识管理、翻译过程管理、翻译业务管理、译员管理等，这些管理模块的实现是在一个统一的网络平台上实现的。因此，对于一个翻译公司而言，实现网络翻译服务平台是该公司在以后竞争激烈的翻译产业中立足的根本。

### 1.2 “人助机译”与“机助人译”

“机器翻译”和“计算机辅助翻译”这两个术语已经被广为接受和普遍使用。一般

来说,前者主要指的是由计算机进行翻译,而后者是由人工在计算机辅助下进行翻译。也就是说,机器翻译是以计算机为主导实现的翻译形式(machine-dominant translation),计算机辅助翻译是以人为主导实现的翻译形式(human-dominant translation)。实际上,目前的机器翻译研究和已经实现的机器翻译系统都脱离不了“辅助(aids)”,如“前编辑(pre-editing)”“后编辑(post-editing)”“句法调序(syntax-based reordering)”等。对这两种都借助辅助手段的翻译技术的区分,我们集中在两个术语——“机助人译”和“人助机译”上,前者是以人为主导实现的翻译形式,后者是以机器为主导实现的翻译形式。

### 1.2.1 人和机器在翻译活动中的分工

无论是“人助机译”还是“机助人译”,参与翻译活动的主体总有两种(并不一定是“两个”)——人(语言学家、技术人员或译员)和机器(计算机)。然而,这两种主体在“人助机译”和“机助人译”中的地位 and 分工是完全不同的。

在人助机译系统中,计算机是最终实现翻译的主体,语言学家和技术人员是为计算机服务的,或者说,是“辅助”计算机的。在机器翻译研究过程中,出现了两种研究范式,一种是基于规则(rule-based)的机器翻译,另一种是基于经验的(empirical-based)机器翻译。基于规则的机器翻译系统的实现可以是直接翻译的,也可以是利用中间语的,但其本质上都是根据总结的规则将一种树形表述(源语句法分析)转换为另一种树形表述(目标语生成),因此,这种实现方式又被称为“基于转换的机器翻译”。在这个过程中,转换是由计算机自动实现的,但提取规则(语法规则、语义规则)、源语句法分析和生成目标语都离不开语言学家和技术人员的辅助。转换过程中常用的语言规则包括:切分规则、标注规则、句法分析规则、语义分析规则、结构转换规则、词语转换规则、结构生成规则、词语生成规则等(张霄军,2010:83-84)。虽然每一套规则最终都可以通过机器学习的方式自动提取,但训练机器的语料和规则却是人工标注和总结的。对源语的句法分析可以由句法分析器自动完成,但自动分析的结果并不完全可靠,需要人工进行校对。目标语生成以后翻译过程似乎就完成了,但这样的译文是“不忍卒读”的。这种机器翻译的译文只是“初稿”,需要校对者进行校对,即“后编辑”。但机器译文的校对和人工译文的校对是完全不同的,正如 Hutchins (2003:509)所言,这样的机器翻译“不适合专业译员,因为他们不喜欢修改计算机程序所犯下

的让人恼火的“低级”错误”。

到了上世纪 80 年代，随着平行语料库的兴起，基于经验的机器翻译研究范式被逐步建立起来了。在这种研究范式下，基于实例的（example-based）机器翻译和基于统计的（statistical-based）机器翻译方法都得到了广泛应用。基于实例的机器翻译方法的基本思想是：利用过去已经翻译过的语料，采用匹配的方法来实现翻译。它假设与例句相同的源语部分对应相同的翻译结果，当出现与例句不相同的源语片段时就需要重组，即在别的例句中寻找相同的源语片段。从原理上来看，基于实例的机器翻译已经脱离了语言（语法）的限制，而彻底脱离语言和语法桎梏的机器翻译方法是统计机器翻译。基于统计的机器翻译方法的基本观点认为，一种语言中的任何一个句子都有可能是另一语言中的某个句子的译文，只不过可能性有大小。那么翻译的任务就是：利用概率知识找出其最大的可能性。具体做法为：在一个理想的规模足够大的平行语料库中统计出源语片段（单词、短语、句子）与目标语片段的同现概率（即在有片段 A 出现的语境中，片段 B 出现的概率）。当输入一段待翻译的源语时，系统根据统计结果选择最大同现概率的译文并输出。基于统计的机器翻译方法不仅是“非语言学的（anti-linguistics）”（Somers,2003:516），看起来也是“非辅助的（anti-aids）”。但实际情况并非如此，无论是基于实例的还是基于统计的机器翻译系统都离不开一个已经标注完备且源语与目标语片段相“对齐（alignment）”的平行语料库，这就既离不开语言学知识，也离不开人工辅助了。而且这样的平行语料库的规模要足够庞大，足以保证输入机器翻译系统的语言片段都能在语料库中检索到，否则就会因没有足够数据支持而出现统计上的错误，即“数据稀疏（data sparse）”（Koehn,2010:182），从而影响到高质量译文的输出。美国统计机器翻译专家 Och 曾经模仿伟大的希腊科学家阿基米德（Archimedes）的口吻放言：“只要给我充分的平行语言数据，那么，对于任何的两种语言，我就可以在几小时之内构造出一个机器翻译系统”（Mankin,2003），但即使如 Google 这样的多语在线搜索引擎，也无法提供一个如此大规模的平行语料库，更何况这样的语料库还要进行标注和对齐，那实现的难度之大就可想而知了。

在机助人译系统中，人是最终实现翻译的主体，计算机与信息技术是为译员和翻译项目参与者服务的，或者说，是“辅助”人的。在计算机辅助翻译系统中，



常用的计算机技术和系统包括文本处理技术、术语管理系统和翻译记忆系统等。要让计算机辅助我们做翻译，首先就得让计算机“阅读”我们的文本，纸质文本计算机是看不懂的，所以需要文本处理技术。应用文本处理技术不仅可以把纸质文本变成机器可读的电子文本，而且可以利用计算机的统计功能来统计词频和字数，这对翻译项目管理非常有用。对于一个合作的翻译项目而言，术语翻译的一致性非常重要，因此，参与该项目的译员需要一个统一的术语管理系统。我们可以利用计算机数据库方便地存储和检索等功能为一个项目创建术语平台，各个译员在翻译过程中可以随时调用该数据库，而且能根据需要进行术语词条的增删。同时，除了词典中的专业词汇，计算机还可以根据文本内容自动抽取和识别新术语。翻译记忆系统是计算机辅助翻译的核心内容，其工作原理类似于上述基于实例的机器翻译的工作原理：二者都是利用已经翻译过的语料，采用匹配的方法实现新的翻译，从而达到语料的复用。二者的区别在于其实现形式：如果输入的句子与语料库中的某个句子完全匹配（完全相同），基于实例的机器翻译系统会自动将其翻译为语料库中的该句对应的译文，而翻译记忆系统则会提示译员完全匹配信息，等待译员确定该句译文。如果输入的句子与语料库中的某个句子不完全匹配，基于实例的机器翻译系统就会选择匹配片段的译文后而在语料库中寻找与其余片段相匹配的句子，而翻译记忆系统则会告诉译员该句与语料库的某个句子匹配的相似度，等待译员人工翻译不匹配的句子片段。例如在平行语料库（翻译记忆库）中已经存储了例句 1.1、例句 1.2 及其各自译文，而系统输入的待翻译的原文为例句 1.3：

例句 1.1 He bought a book on international politics.他买了一本有关国际政治的书。

例句 1.2 I read a book on international politics.我读了一本有关国际政治的书。

例句 1.3 I bought a book on Monday.

基于实例的机器翻译系统的处理过程是：系统首先判断例句 1.3 与语料库中的任何原文都不是完全匹配；接下来它发现第一个语言片段“T”与例句 1.2 的语言片段“T”完全匹配，于是优先翻译该语言片段——选择了例句 1.2 的“T”相对应的译文“我”；然后重组发现语言片段“bought a book on”与例句 1.1 的语言片段“bought a