

# 前 言

佛山市雪人计算机有限公司是一家专注于软件开发及销售的高科技公司，公司一直以开发和推广自主知识产权的软件为目标，致力于行业软件的开发。公司坚持走自主创新、稳步发展之路，以“做国内最好用的行业软件”为目标，以产品自身的品质、先进的技术和良好的服务取信于用户。

雪人计算机辅助翻译软件(Snowman Computer Assisted Translation Software——缩写为：雪人 CAT)是一种充分利用计算机的超强计算能力、记忆能力和人的创造能力相结合的人机互动的辅助翻译软件，由译员把握翻译质量，计算机提供辅助。它能够辅助译员优质、高效、轻松地完成翻译工作，帮助企业及个人充分利用资源、降低成本、成倍提高工作效率。适用于需要精确翻译的机构和个人。

雪人CAT软件特色：**1. 简单易用、速度快：**瞬间完成在百万级的记忆库中整句匹配或片段搜索；**2. 准确、快速的双语对齐句级对齐：**轻松创建大型记忆库。**3. 实时预览技术：**译文预览随译随见；还支持预览原文和译文、原文混合预览。**4. 自动替换翻译：**雪人采用EBMT和TM两种技术相结合，实现自动替换翻译。**5. 嵌入在线自动翻译和在线词典：**自动给出“在线自动翻译”的译文供参考；嵌入“Google、EngKoo、youdao、iCIBA、Dict.cn、JuKu”等多个在线词典，切换方便。**6. 用户自定义语法规则翻译，**进一步提高取词和翻译的效率。**7. 质量检查功能：**可以检查术语一致、数字校验、漏译、拼写、错别字、一句多译等等。**8. 预翻译：**利用记忆库、词典、规则词典、本地术语库和在线翻译相结合等方式进行预翻译，快速生成译文。

为了让更多的译员了解雪人CAT软件，我们推出了免费版。免费版带有30万词条的小型词典，支持TXT格式的文本文件，同时支持导入TMX格式的记忆库和XML格式的词典文件，可以胜任实际的翻译工作。

本软件的使用说明书由黄予等人编写，由于时间紧迫，难免有错漏之处，欢迎广大用户指正！请将您的建议或发现的问题发送到xueren@gcys.cn邮箱中，谢谢！

另外，雪人CAT软件自2009年11月份推出以来，收到众多免费版、标准版用户提出的宝贵意见，雪人公司在此表示衷心的感谢！我们会不断地完善软件，使它越来越好用！

佛山市雪人计算机有限公司

电话：0757-83216691

网站：<http://www.gcys.cn>

邮箱：[xueren@gcys.cn](mailto:xueren@gcys.cn)

(网站上有 CAT 的教学视频)

# 目 录

前 言.....	1
目 录.....	2
雪人CAT软件特点.....	8
1. 界面简洁、易学易用.....	8
2. 运行稳定、快速.....	8
3. 实时预览技术.....	8
4. 嵌入在线自动翻译.....	10
5. 嵌入多个在线词典.....	10
6. EBMT技术，实现准确的自动替换翻译.....	11
7. 同时支持表格和点击输入译文两种界面.....	12
8. 快速响应的片段搜索.....	14
9. 智能处理词形变化.....	14
10. 导出双语对照文档.....	14
11. 智能的屏幕取词功能.....	15
12. 利用剪切板功能翻译各种格式的文件.....	16
13. 强大的双语对齐工具帮助译员快速创建大型记忆库.....	16
14. 词语的整批替换.....	16
15. 词频统计.....	16
16. 自定义语法规则翻译.....	16
17. 预翻译.....	16
第一章 常用热键.....	17
第二章 快速入门.....	19
第一节 软件的安装.....	19
第二节 快速入门.....	21
一、新建项目及引用词典、记忆库.....	21
二、在对照模式中翻译.....	25
三、在单句模式中翻译.....	29
四、审稿、校稿.....	33
五、项目词典和记忆库的维护.....	35
六、导出译文或双语对照.....	35
第三章 软件常用功能介绍.....	37

第一节 句子的拆分与合并.....	37
第二节 定义术语和定义新词.....	37
第三节 自动调整取词顺序.....	39
第四节 自己建词库(项目词典).....	41
第五节 自己建记忆库(项目记忆库).....	46
第六节 其他功能.....	51
一、单行对比清晰显示两句的异同.....	51
二、鼠标划选方式.....	51
三、片段搜索的显示模式.....	53
四、常用字自定义分页.....	55
五、方便的查找功能.....	57
六、自动存盘.....	57
七、手工分词.....	58
八、常用符号工具栏.....	59
九、识别大写词语选项.....	59
十、支持使用代理服务器.....	61
第四章 翻译各种文件格式.....	62
第一节 翻译WORD文档.....	62
一、对图文并茂WORD文档的支持.....	62
二、对表格的支持.....	66
三、对复杂数学公式的支持.....	68
四、对复杂化学式的支持.....	70
五、隐藏预览区.....	71
六、复制原文及组合译文.....	71
第二节 翻译PPT文档.....	72
第三节 翻译EXCEL文档.....	74
第四节 利用剪切板翻译各种格式的文档.....	77
一、翻译网页.....	77
二、翻译PDF文件.....	78
第五节 导出句子双语对照和Unclean格式文件.....	80
一、导出WORD格式的句子双语对照.....	80
二、导出EXCEL格式的句子双语对照.....	81
三、导出PPT格式的句子双语对照.....	82

四、导出Unclean格式的Word双语文件 .....	83
第五章 网络资源的整合运用 .....	84
第一节 嵌入多个在线词典 .....	84
第二节 嵌入在线自动翻译 .....	88
第三节 用户术语库与在线自动翻译的结合 .....	95
第四节 在线搜索 .....	97
第六章 预翻译 .....	100
一、利用记忆库及规则词典预翻译 .....	100
二、利用在线翻译进行预翻译 .....	103
三、检查记忆库多重匹配 .....	106
第七章 自定义断句规则 .....	107
第八章 EBMT技术实现自动替换翻译 .....	113
一、实现一般的替换翻译 .....	113
二、实现增词替换翻译 .....	114
三、实现减词替换翻译 .....	116
四、方便的短语提取 .....	116
五、创建高质量的模板型记忆库 .....	118
第九章 词语的整批替换 .....	121
第一节 译文的整批替换 .....	121
第二节 正则表达式 .....	123
一、正则表达式理论知识 .....	123
二、正在表达式实际应用 .....	124
1. 将句尾的“.”号全部替换成“。”号 .....	124
2. 搜索多个词 .....	125
3. 精确查找单词 .....	125
4. 对间隔短语的搜索 .....	127
5. 中文搜索 .....	128
三、自动记录查找条件 .....	129
四、全字匹配 .....	131
第三节 对查找的结果直接定位到原处 .....	132
第四节 设置默认的替换内容 .....	133
第十章 词频统计（词语提取） .....	134
一、从待译文章中提取词语 .....	134



二、从记忆库中提取词语 .....	138
三、查看长句.....	139
四、词频统计热点词汇 .....	140
五、词频统计的词语搭配分析 .....	140
第十一章 双语对齐创建大型记忆库 .....	143
一、中英文分开的对齐 .....	145
二、中英混排的对齐 .....	148
三、直接导入其他CAT软件译稿制作记忆库 .....	150
四、自动对齐后的编辑 .....	153
五、同步修改双语对齐 .....	154
六、导出记忆库 .....	156
七、导出双语对照文件 .....	156
八、记忆库的直接转换 .....	159
九、双语对齐的不合并句子选项 .....	159
第十二章 独立的记忆库搜索 .....	163
第十三章 自定义语法规则翻译 .....	165
一、什么是类别 .....	165
二、规则定义的常用符号: .....	165
三、优先级 .....	174
四、多重嵌套 .....	174
五、语法规则的共享 .....	174
六、临时规则 .....	174
七、用更简洁的方法定义规则词典 .....	175
八、在EXCEL中编辑语法规则词典 .....	177
第十四章 质量检查 .....	178
第十五章 利用雪人CAT进行审校 .....	182
第十六章 网络协同翻译平台 .....	185
第一节 客户端与服务器连接 .....	185
第二节 实时共享记忆库、术语库 .....	189
第三节 审校与翻译同步 .....	193
第十七章 版本区别及配置要求 .....	198
一、免费版和标准版的功能区别 .....	198
二、32 位和 64 位的区别 .....	199

三、雪人CAT对机器的配置要求 .....	199
第十八章 常见问题解答 .....	200
1. 我该选购 32 位版本还是 64 位? .....	200
2. 如何开始翻译工作?.....	200
3. 我用双语对齐做的记忆库不能导入怎么办? .....	200
4. 词典、记忆库的添加引用与导入有什么不同? .....	200
5. 已经添加了记忆库，怎么在“项目记忆库”还是空白? .....	201
6. 雪人可以导入哪些格式的词典和记忆库? .....	201
7. 如何制作EXCEL格式的句库或词典导入到雪人中? .....	201
8. 我有很多翻译资源，如何将它们转入到雪人中来用呢? .....	201
9. 记忆库大了，软件使用起来会不会感觉较慢? .....	202
10. 记忆库的最小匹配度与片段搜索的区别? .....	202
11. 雪人与trados等的兼容性如何? .....	202
12. 定义术语和定义新词的区别?.....	202
13. 如何进行句子的拆分与合并? .....	202
14. 需要拷贝哪些文件就可以将工作移到其他机器上? .....	202
15. 免费版词汇量有上限吗?.....	203
16. 免费版没有词典和记忆库合并功能，该如何合并呢? .....	203
17. 如何维护雪人的项目记忆库、项目词典? .....	203
18. 雪人正版软件能在几台机器上用?.....	203
19. 雪人正版软件的升级是免费的吗? .....	203
21. 为何在Vista或Windows 7 下无法保存设置状态? .....	204
22. 为何免费版本的“在线词典”显示为空白? .....	204
23. 免费版直接双击STP文件为何不能打开? 该如何关联? .....	204
24. 如何设置操作界面的背景色及字体大小? .....	204
25. 导入、导出WORD文档时怎么软件好像没有反映? .....	206
26. 导出WORD文档时出错，怎么办? .....	206
27. 导入EXCEL文件时出错怎么办? .....	206
28. 导出EXCEL时为什么有时一直无法完成? .....	206
29. 使用Office2007 怎样生成Office2003 兼容的文档? .....	207
30. PDF转WORD一句插入了其他内容而被拆成多句的翻译? .....	207
31. 异地如何进行多人协同翻译? .....	207
32. 为什么出现保存的文件不能打开? .....	207

33. 导入PPT文件出错时怎么办? .....	207
34. 为什么PPT组合内容没有导入? .....	207
35. 为什么WORD中带批注不能导出? .....	207
36. 为什么WORD的文本框内容没有导入? .....	208
37. 为什么WORD2010 Starter版不能正确导入? .....	208
38. 为什么“在线自动翻译”不能使用? .....	208
39. 导入原文时，显示转换器的问题，是怎么回事? .....	208
40. 如何修改所定义的术语、词语? .....	208
41. 如何快速处理非译成分（或快速复制原文）? .....	209
温馨提示 .....	209

# 雪人 CAT 软件特点

## 1. 界面简洁、易学易用

雪人简化了记忆翻译的概念，将繁杂的功能整合在高效易用的界面中，使译员可以在很短的时间内掌握 CAT 软件的使用方法。

## 2. 运行稳定、快速

- (1). 轻松支持百万级的翻译记忆库。在百万级以上的记忆库中搜索、匹配无须任何等待。
- (2). 一分钟导入 15 万 TMX 格式的句对到记忆库，并且在导入时已经删除了重复的句子和统计出实际例句数量，确保译员自建的记忆库中同样的句子不会出现重复。
- (3). 一流的软件稳定性，确保您的工作成果。

## 3. 实时预览技术

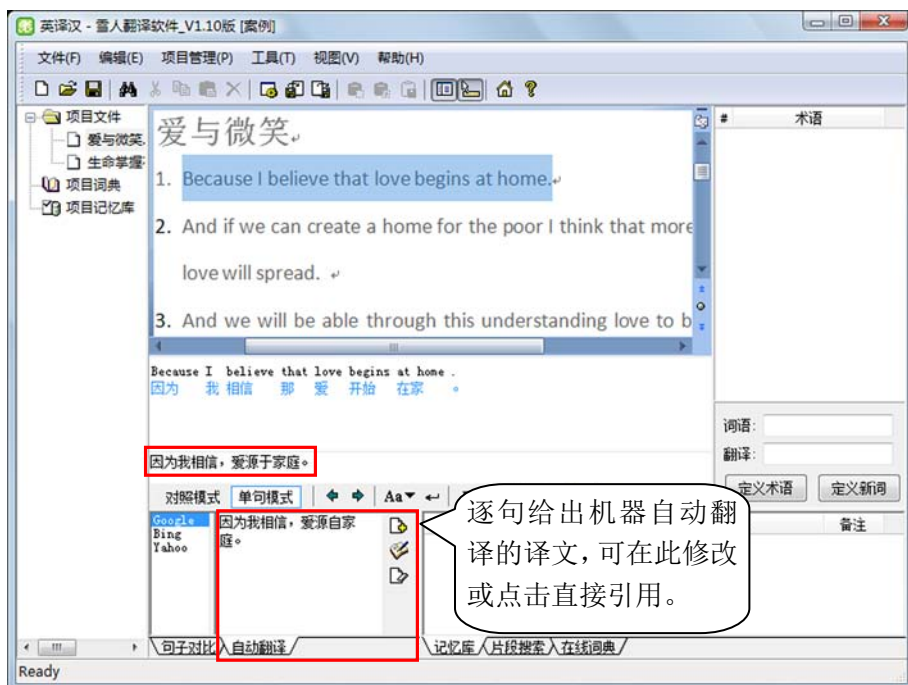
译文预览随译随见，即时自动更新，无需译员手动刷新；还支持预览原文和译文、原文混合预览，非常直观、方便。使译员可以更好地把握文章的整体性，对样式复杂的文档、数学公式、化学分子式、化学方程式等文档的翻译尤显方便。如下图：





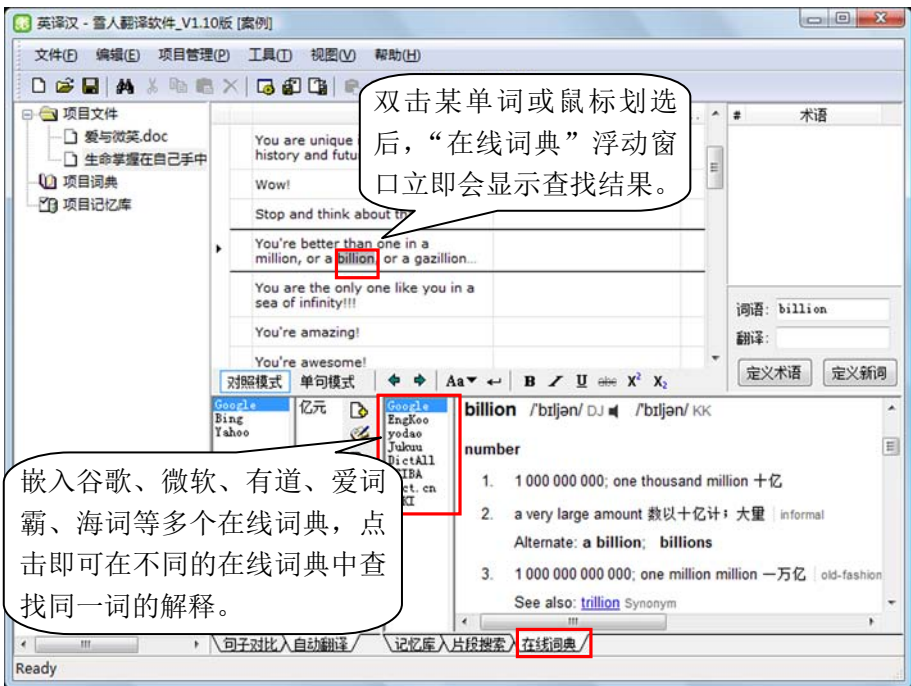
## 4. 嵌入在线自动翻译

雪人在软件中嵌入“Google、Bing、Yahoo、Youdao、Systrant ”等多个自动翻译引擎，逐句自动给出“在线自动翻译”的译文供参考，点击即可引用或在此基础上修改。



## 5. 嵌入多个在线词典

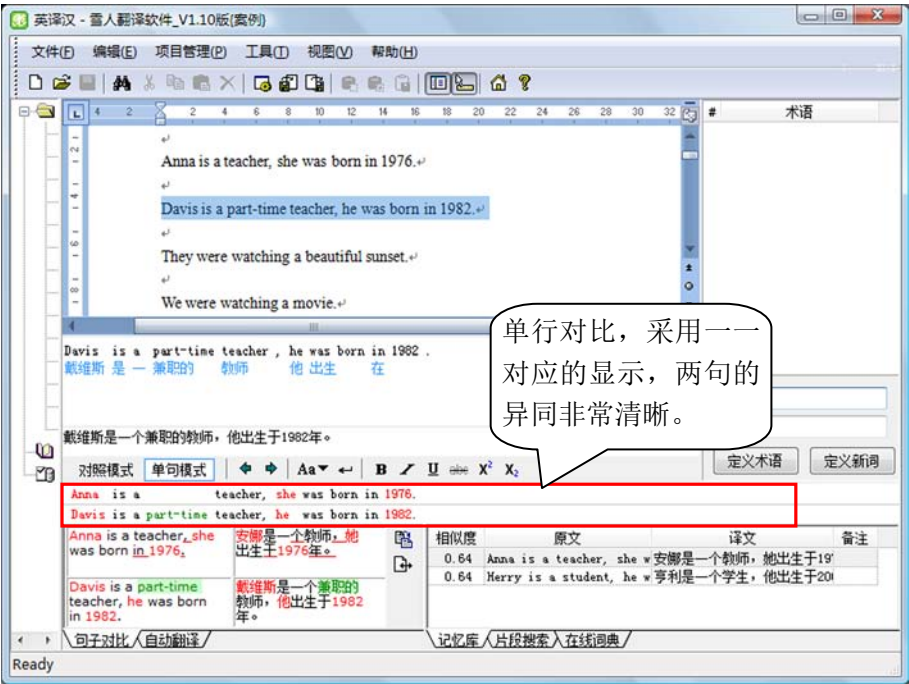
雪人系统自带词典非常丰富，标准版有 1200 万、涵盖 60 多个专业领域的词库（免费版有 30 万的词汇）。同时，还很好地利用了在线词典的资源为译员服务，雪人在软件中嵌入“Google、EngKoo、youdao、Jukuu、iCIBA、Dict.cn、DictAll、CNKI、FastDict、CiB0.cn”等 10 个在线词典，译员只要在原文中双击或用鼠标划选某词语后，立即显示查找的结果，点击即可切换到其他在线词典查看。



## 6. EBMT 技术, 实现准确的自动替换翻译

雪人软件采用 EBMT(Example-Based Machine Translation 基于实例的机器翻译)和 TM (Translation Memory 翻译记忆) 两种技术相结合, 充分发挥翻译记忆库的作用。目前的 CAT 软件主要采用的是 TM 技术, 而单纯的 TM 技术存在着精确匹配率不高, 模糊匹配时产生译文质量较差等的缺点。雪人 CAT 采用先进的 EBMT 技术, 通过类比、推理、学习等方法, 可以根据记忆库中已有的例句, 自动替换翻译出其他相识的句子。

例如: “He is a student. 他是一个学生。” 与 “America is a developed country. ” 这两句看似相差甚远, 一般的 CAT 软件根本就不认为它们之间有联系。雪人却能以记忆库中的 “He is a student. 他是一个学生。” 这一句为模板, 将 “America is a developed country. ” 准确翻译为 “美国是一个发达的国家。此外, 雪人还能以记忆库中的例句为模板, 进行增词、减词的替换翻译, 并以不同颜色清晰地标注两句之间的异同。如下图:

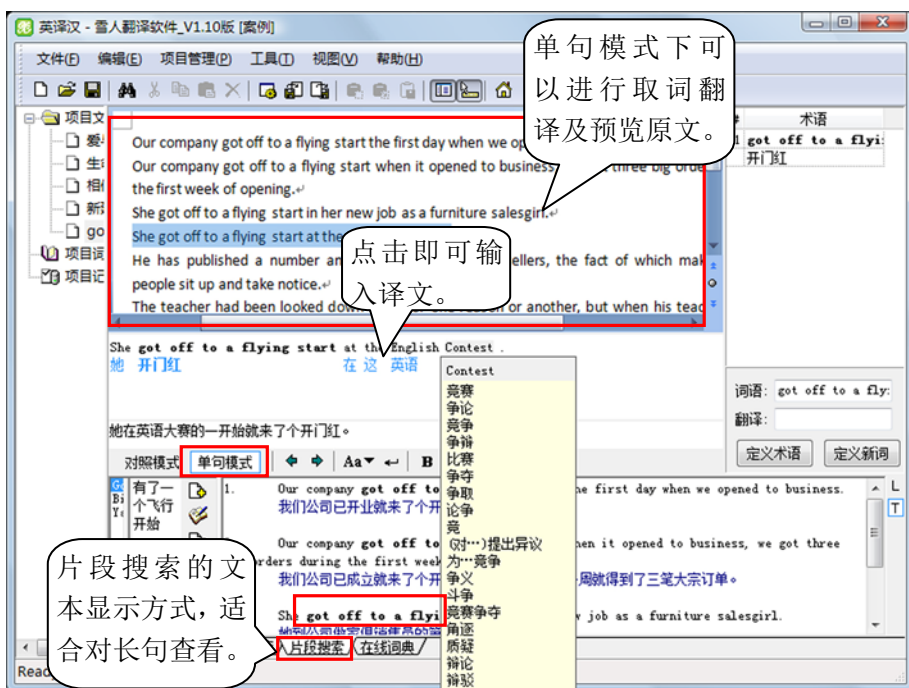
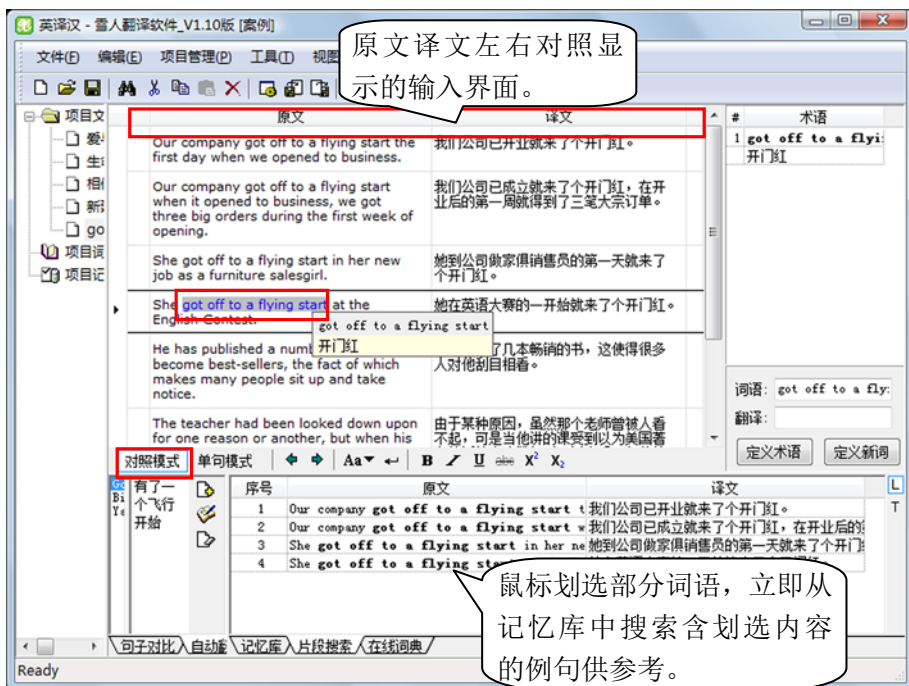


7. 同时支持表格和点击输入译文两种界面

支持左右表格排列的“对照模式”和“单句模式”两种输入模式，用户可以根据个人习惯采用，两种输入模式可以随时切换。“对照模式”输入界面（如下图）：原文、译文以左右对照的表格方式排列，这种界面简洁高效，条理清晰，对表格和审稿、校稿尤其方便。

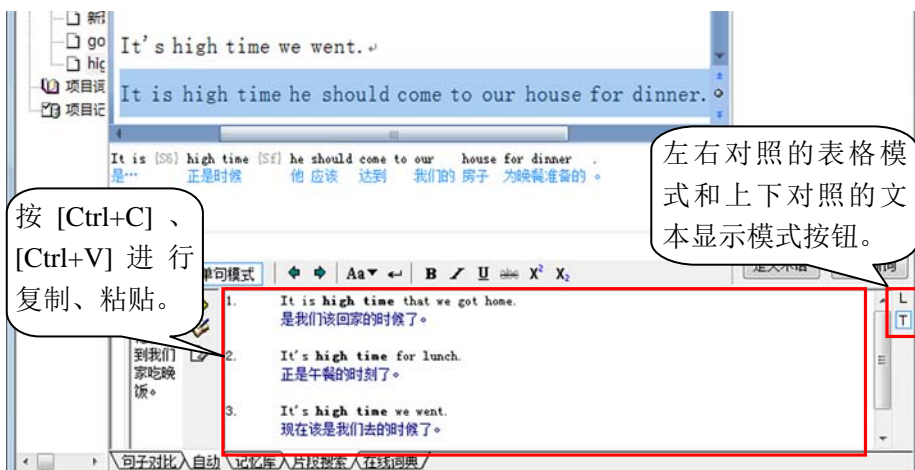
“单句模式”输入界面（如下图）：可以预览原文、译文和进行自动的取词翻译，鼠标点击即可输入译文，减少击键次数。这种界面直观、整体性强，对样式复杂的文档和公式、化学式等尤其适用。





## 8. 快速响应的片段搜索

由于能完全精确匹配的例句并不多，从记忆库中对部分词语、短语给出译文参考，这对于译员来说也是相当有价值的。雪人 CAT 软件可以很好地实现这一功能，只需在原文用鼠标划选某片段后，在下方的“片段搜索”窗口立即显示出搜索结果，而且可以将查找的结果复制到译文中或定义新词中（按快捷键 **【Ctrl+C】**、**【Ctrl+V】** 进行复制、粘贴）。

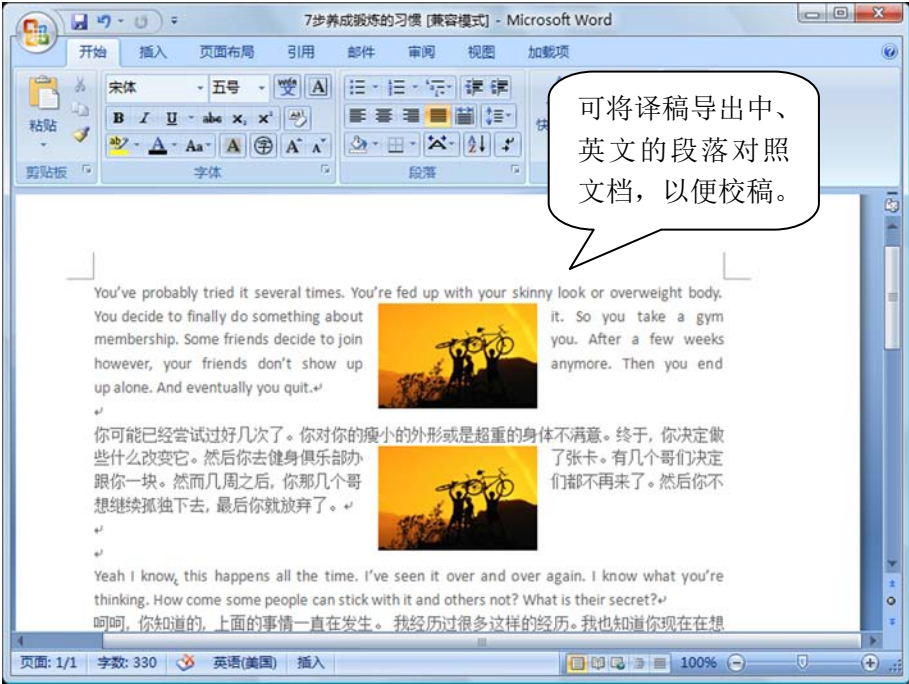


## 9. 智能处理词形变化

不会因为某词语的时态变化、单复数的不同而影响搜索、匹配的结果（免费版不支持这个功能），充分发挥记忆库的作用。

## 10. 导出双语对照文档

为了方便译员和翻译公司之间进行文件交流，雪人可以将译稿输出中英文的段落对照和网页格式的双语对照文档，方便校稿。段落格式即一段原文、一段译文。网页格式以句为单位，一句原文、一句译文。



11. 智能的屏幕取词功能

目前大多数译员是一个 CAT 软件加一个词典软件协同工作，而且还存在有些 CAT 软件和词典软件冲突的问题，查到的词语也只能再次录入或复制。在雪人 CAT 中当鼠标指向某词或短语停留片刻，会弹出屏幕取词窗口，单击内容后可以输入到译文中，极大地减少了译员键盘输入的工作量。

雪人 CAT 软件另一方便之处是译员在翻译过程中定义的新词、术语和短句，可以立即在屏幕取词中应用。例如在上一句翻译时将“set the world on fire ” 翻译为“轰动全世界”，并将它保存到了词典，当下一句再次碰到该短语时，立即会弹出该译文，单击它即可输入。这是利用其他 CAT 软件和词典软件协同工作时比较难实现的。



## 12. 利用剪切板功能翻译各种格式的文件

雪人具有非常强大的剪切板功能，通过它，可以翻译网页、PDF、PPT 等等一切可以复制粘贴的文档。

## 13. 强大的双语对齐工具帮助译员快速创建大型记忆库

无论是网上看到的双语文章，还是译者自己以前翻译好的文章；无论是一篇中文、一篇英文分开的，还是中英文混排的文章，雪人翻译辅助软件都可以自动一句一句地帮你对齐好，而且还可以直接读入其他 CAT 软件的译稿制作成自己的记忆库，可以帮你轻松创建大型的记忆库！

## 14. 词语的整批替换

功能强大的“查找和替换”功能，方便对译文、词典、记忆库进行整批的修改替换。支持“正则表达式”，查找方法灵活，目标结果准确，想找什么都能找到，查到结果后可以对它们进行整批的替换，修改起来十分方便。

## 15. 词频统计

若在翻译前预先统计出待译文章中那些重复率高的词语，对它们进行预先的准确翻译，这将给后续的翻译工作带来极大的方便。此外，利用“词频统计”功能还可以从记忆库中提取高频词语，将它们添加到词典中。

## 16. 自定义语法规则翻译

无论是英语还是汉语，都有一定的语法规则可循，用户可以在规则词典中预先将这些规律定义成语法规则，这样可以进一步提高取词和翻译的准确性。例如“be satisfied with...，对.....感到满意”，若将它定义到规则词典中，则以后若遇到“be satisfied with you”就可以准确翻译出：“对**你**感到满意”，若遇到“be satisfied the result”就可以将它准确译成：“对**这个结果**感到满意”。

## 17. 预翻译

利用记忆库、词典、规则词典、本地术语库和在线翻译相结合等方式进行预翻译，快速生成译文。

# 第一章 常用热键

【F2】——浏览与修改状态切换，浏览状态下光标以“整行的表格”（以蓝色底色显示）为单位进行上下移动，移动速度快，在项目词典或记忆库窗口中，浏览状态下按【Ctrl+Delete】可以整行删除词典或记忆库中的句子。修改状态下光标以表格“单元内的行”（以白色底色显示）为单位进行上下移动，移动速度慢，在项目词典或记忆库窗口中，修改状态下可以对词典或记忆库的原文、译文进行再次编辑、修改。

【F3】——在利用查找功能进行查找时，按【F3】可以查找下一目标。

【F4】——在第一列对有问题的译文行以“×”在表示，以示引起注意。修改好后，再按一次【F4】可以将×设置为黄色的修改图标，表示已经做了修改。

【F5】——在项目词典中，将某一词语设为术语，并以粗体显示。翻译时，按【F5】可复制样式码。

【F6】——复制参考译文。在编辑状态下，将下方从记忆库中搜索到的匹配参考译文复制到当前译文区，并对不同的地方进行了自动替换翻译。【Ctrl+F6】不使用自动翻译替换。

【F7】——复制“在线自动翻译”窗口的译文到译文区中。

【F8】——将原文直接复制到译文区。

【F9】——对某行核对正确，按下【F9】将在该行出现“√”，表示审核通过，再按撤销。

【F11】——用于快速切换是否隐藏屏幕左侧的“项目列表”窗口。

【F12】——用于在“单句模式”快速切换是否隐藏“预览窗口”，以获取更大的编辑区域，对长句的翻译尤显方便。

【Ctrl+N】——直接跳到有“×”标志的行进行修改。

【Ctrl+D】——定义术语

【Ctrl+W】——定义新词

【Ctrl+Z】——在编辑状态下，可以后退、撤销刚才所进行的修改。

【Ctrl+Y】——在编辑状态下，可以恢复刚才所进行的撤销。

【Ctrl+↑】——翻译上一句。

【Ctrl+↓】——翻译下一句。

【Ctrl+T】——在单句模式下，全选后按【Ctrl+T】可以将“{}”里面的内容变成样式代码，样式码会以灰色显示。否则“{}”里面的内容不认为是样式代码。

【Ctrl+R】——查找和替换。

【Alt+Ctrl+P】——在单句模式下，切换到页面视图。

【Alt+Ctrl+O】——在单句模式下，切换到大纲视图。

【Alt+Ctrl+N】——在单句模式下，切换到普通视图。

【Ctrl+\\——】在主档文档和子文档之间移动。

【Ctrl+鼠标滑轮向前】——在单句模式下，放大。

【Ctrl+鼠标滑轮向后】——在单句模式下，缩小。

【Ctrl+ Delete】——在浏览状态下，删除整行记录（在对词典或记忆库进行维护时适用）。

【Ctrl+ 】——跳转到上一句或下一句未翻译的句子。

【Ctrl+回车键】——跳转到下一句未翻译的句子。

【Ctrl+L】——大写转小写快捷键。鼠标划选词语后单击右键，会弹出“词语定义”窗口，光标在输入框中，按【Ctrl+L】可以将大写字母转换为小写。

**拆分句子**——在“对照模式”中，光标在原文需要断句的位置处按【Enter】回车键，即可拆分成两句。

**合并句子**——在“对照模式”中，光标在一行的末尾处时，按【Delete】删除键，即可将下一句合并到上一句中（注：“单句模式”中，不支持句子的拆分与合并，可以切换到“对照模式”中进行拆分、合并，然后再切换回去，在两种输入模式中进行切换时，软件会自动定位当前翻译的句子）。

## 第二章 快速入门

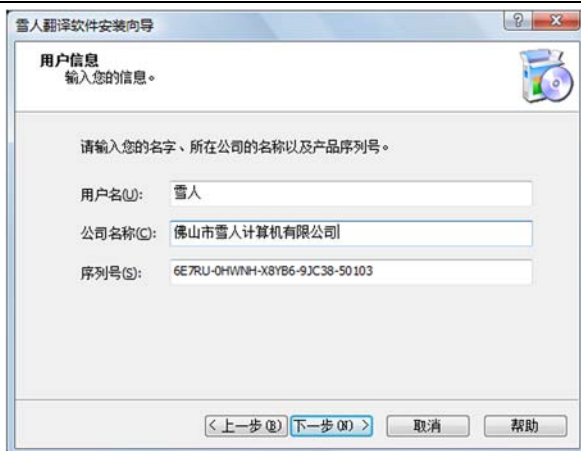
### 第一节 软件的安装

在安装雪人翻译软件之前，请先检查你的计算机是否安装了 WINDOWS XP 及以下的操作系统，是否能正常运行？软件的安装很简单，具体步骤如下：

1. 关闭计算机，将软件的 USB 加密锁插入计算机的 USB 接口上，然后启动计算机。建议：装、拆加密锁时最好在计算机关闭状态下进行，带电装拆易损坏加密锁。
2. 将光盘放入光驱中，稍等片刻。双击桌面上**我的电脑**，然后双击你的**光驱**所在的盘，双击光盘中的 SETUP 安装图标启动安装向导，如下图，点击**下一步**按钮后会提示你需要安装软件还是升级软件，升级程序与安装程序是同一个文件，既可以用于安装，又可用于升级。如果需要安装软件，则选择“重新安装软件”选项，如果已经安装了软件需要升级的话，则选择“升级软件”选项。然后点击**下一步**按钮。



3. 在接下来的“许可协议”向导窗口中选择“我同意许可证协议中的条款”选项，**下一步**按钮才可以点击，否则，**下一步**按钮将显示为灰色不可按，然后单击**下一步**。



4. 输入用户名、公司名称、序列号，**注意：这三项为必填项，而且序列号应该在英文状态下输入，不能在汉字状态下输入，否则下一步会为灰色，不可按。**在**序列号**栏中必须输入正确的序列号，对于同一套软件最多允许安装在三台不同的电脑上，但同时只能在一台电脑上使用，即：使用时必须插上加密锁。加密锁最多只能保存三台不同电脑的安裝记录信息，如果多于三台，则其中一台机器的安装信息将会被覆盖。当你将加密锁插回这台电脑时，需要重新安装后才能使用。为了方便用户重装，加密锁第一次正确安装后会记录安装的序列号，下次安装时会自动填写该序列号，不用再重新输入。正确输入用户名、公司名称、序列号后，单击**下一步**按钮。
5. 选择软件的安装路径，软件的默认安装路径是 c:\Program Files\SCAT\，如果 C 盘空间不大，建议你安装在其他的盘上，如 D 盘，只需将 C 该成 D 即可，当然，你也可以单击旁边的 **...** 按钮来选择安装路径。安装路径的选择确定你将来所有文件的存放目录，如果你的安装路径是：d:\Program Files\SCAT\，则默认状态下，所有的文件将会被存放在该目录下面。单击的**下一步**按钮，将进行软件安装。
6. 正确安装后单击**完成**按钮，关闭安装向导窗口，完成雪人软件的安装工作。

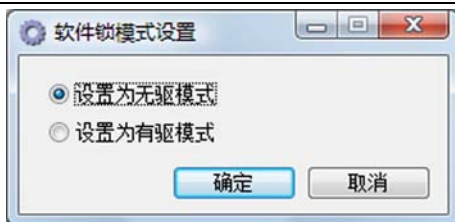
## 安装注意事项：

对于操作系统是 64 位的机器，首先必须将软件加密锁设为“**有驱模式**”并安装软件狗的驱动程序，这样 WINDOWS 才能检测到加密锁的存在（32 位的操作系统可以直接安装无需进行设置）。具体操作方法如下：

1. 打开安装盘中的 TOOL 文件夹，然后双击“setdevicemode”文件，会弹出“软件锁模式设置”窗口，选择“设置为有驱模式”选项后，点击**确定**按钮，就可以将其设为有驱模式了。





2. 设为有驱模式后，跟着就需要安装软件狗的驱动程序。打开光盘的“drivers”文件夹进行驱动程序的安装。**安装完成后，请拔出软件狗，然后再重新插入，这样就可以开始安装雪人翻译软件了。**



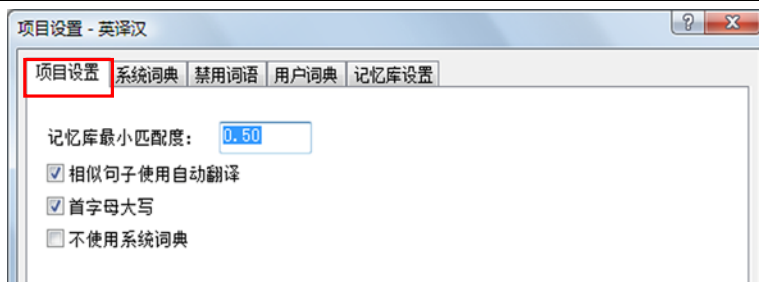
## 第二节 快速入门

### 一、新建项目及引用词典、记忆库

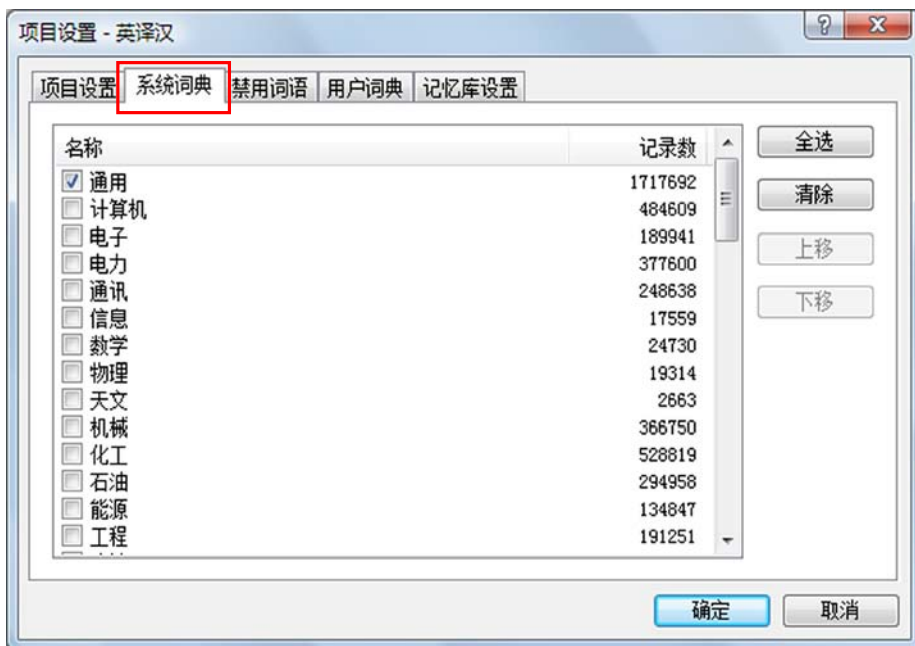
安装好雪人翻译软件后，在桌面上会出现  图标，双击该图标，启动软件，点击  新建按钮，选择项目文件类型是英汉还是汉英，这里以英汉为例。



点击 **确定** 按钮后，接下来会出现项目设置窗口，如下图：其中“记忆库最小匹配度”是指当需要翻译的句子与记忆库中例句的相似程度大于该设定值时，会自动将记忆库中的例子列出供参考，相似程度小于该设定值的例句将不出现，软件默认的最小匹配度为 0.5，用户可以自行设定该值，但必须在 0.5~1.0 之间。若相似度小于 0.5，整句的匹配意义不大，此时应该使用软件另一强大功能——**片段搜索**！“相似句子使用自动翻译”选项是指对相似例句的不同部分，由软件自动做智能的替换翻译。



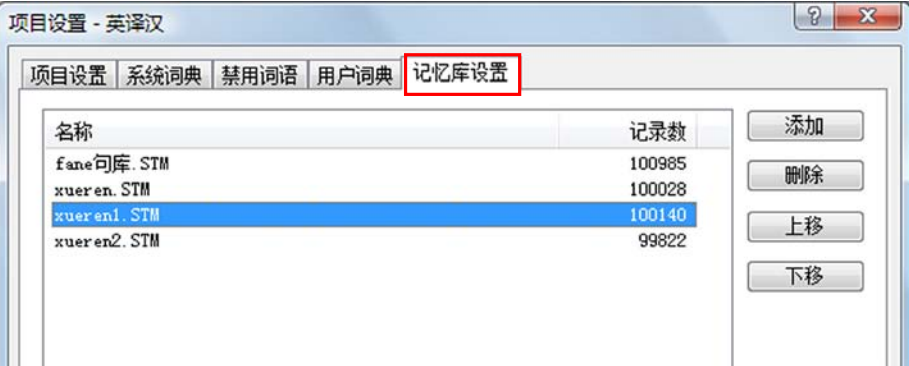
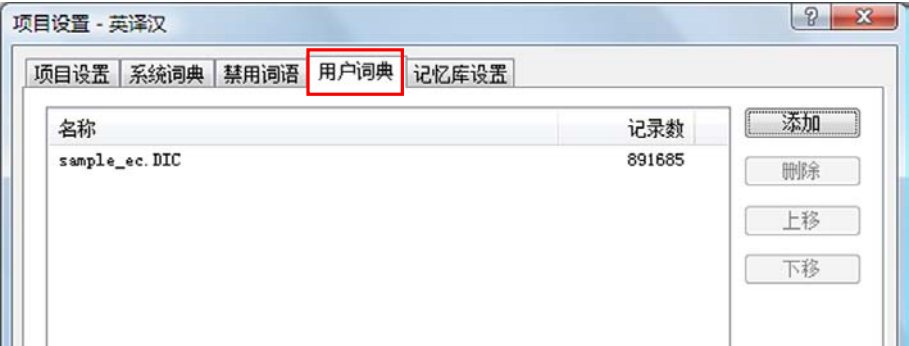
雪人 CAT 软件系统自带数 60 多个专业、千万词的英汉和汉英词典。（免费版系统自带 30 万词的基本英汉、汉英词典）



另外，雪人 CAT 软件有删除系统词条功能，在屏幕取词窗口中使用鼠标右键选择属于系统词典的解释后，可以删除系统的词条。删除功能只在当前项目有效，并且是取消了系统词典中该词条的所有解释。通过删除系统词条，可以更正系统词典中的错误和更灵活地分词。被删除了的系统词条会保存在【项目设置→禁用词语】页签中，如下图：在此可以增加或取消禁用的系统词条项目，点击“清除”按钮可将删除的词条恢复，还可以通过导入、导出实现与其他翻译项目进行交换。

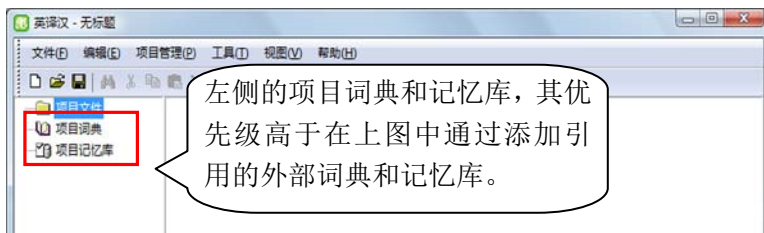


接下来根据翻译项目所属的专业类型进行词典和记忆库的引用设置。如下图：词典和记忆库在下图的排列顺序决定它在这个项目文件应用中的优先顺序，排在上面的优先级别高，以此类推；在此添加的词典或记忆库，软件通过“引用”的方式来使用它，此处添加的为“外部的词典或外部记忆库”。



完成上述设置后，软件会自动生成一个项目文件、一个项目词典和一个项目记忆库（如下图）。右击项目文件，将弹出导入文件右键菜单，导入需要翻译的文件，在一个项目文件中可以导入多份原文，多份原文既可以是相同的文件格式，亦可以不同的格式，例如：TXT、DOC 等等其他文件格式（注：免费版只支持导入 TXT 格式的原文，标准版可以支持导入多


种格式的原文)，项目文件下导入的多份文件共用相同的记忆库、词典。

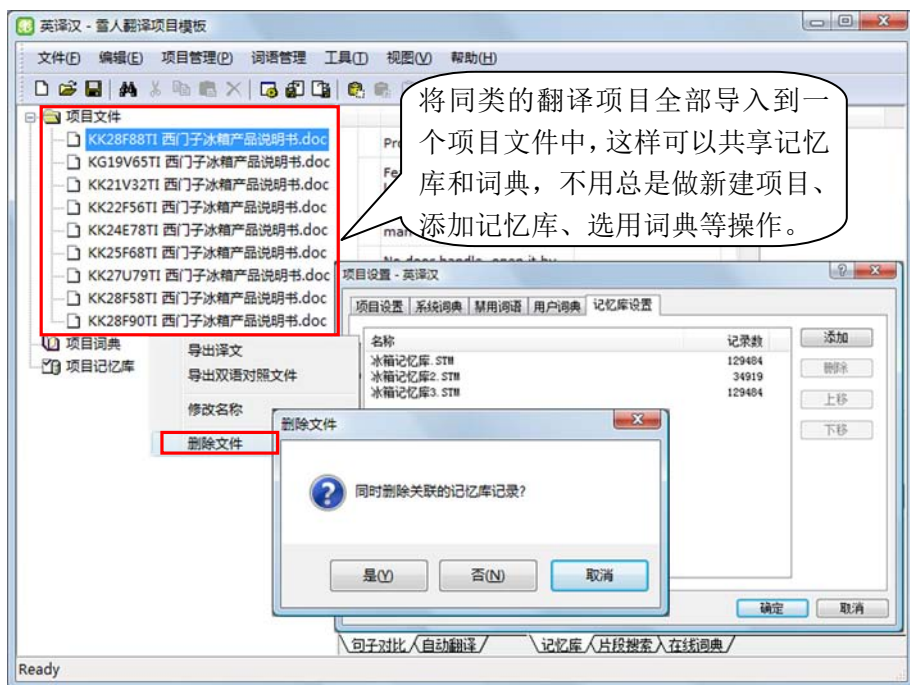


左侧的项目词典和项目记忆库是用于记录本项目在翻译过程中产生的新词汇和新例句，开始为空。译员在翻译过程中添加的新词、术语会被保存到项目词典里，翻译好的译文会被自动保存到项目记忆库中。右击项目词典或项目记忆库，利用【**右击菜单→导入**】功能可以导入其他词典或记忆库为当前翻译项目服务。若整个翻译工作完成后，可以在此对本次翻译积累下来的词典、记忆库进行进一步的整理（【**F2**】浏览与修改状态切换，浏览状态下按【**Ctrl+Delete**】可以整行删除，修改状态下可以对原文、译文进行修改）。然后利用【**右击菜单→导出**】功能或者到【**项目管理→导出**】菜单将积累的成果导出为外部的词典和外部记忆库，以备日后使用。新建项目时在用户词典和记忆库设置那里添加的词典和记忆库属于外部词典和外部记忆库，不会出现在此。

**建议：**尽量通过“添加”引用外部的词典和外部记忆库的方式来为当前翻译项目指定使用的词典和记忆库，而不是将它导入到“项目词典”或“项目记忆库”，因为“添加引用”的方式可以更好地减少项目文件的大小和对系统资源的占用，对配置不高的机器尤其重要。

## 创建翻译项目模板共享记忆库（技巧）：

在需要翻译一个新文件时，其实用户不必每次都去新建一个项目，然后添加记忆库、词典等操作，这样做的繁琐之处是：每个新项目都需要添加记忆库、词典，还要将本次翻译好的记忆库、词典导出、合并，尤其对于不能直接合并记忆库和词典的免费版用户，就显得不便了。有个非常好的解决办法就是：利用雪人在一个“ **项目文件**”下面可以**200**份的待翻译文档的功能，建立一个**翻译项目模板**，当需要翻译新文档时，就打开这个项目模板，将需要翻译的文件导入模板中，例如下图中，将同类的各个冰箱型号的说明书文档全部导入到**翻译项目模板**中。



译好后，可以保留这个文件，若觉得确实没有必要保留，则点击 **项目文件** 下的该文件，然后点击【**右键菜单**→**删除文件**】功能将它删除，删除时会弹出“**是否在删除文件的同时，删除其关联的记忆库记录？**”一般情况下，文件译好后，记忆库肯定是要保留下来的，所以选择“否”，这样该文件则会从左侧的目录树中消失，而它形成的记忆库则保留在 **项目记忆库** 中。如果意外点击了删除记忆库，还可以用放弃存盘以免造成不必要的损失。

利用“翻译项目模板”的好处有：

1. 不用总是进行新建项目、添加记忆库、词典等操作。
2. 不用进行词典、记忆库的导入、导出合并等操作，在同一个项目中不断积累记忆库、词典，给日后所有待翻译的文档共享，这对合并记忆库、词典不方便的免费版用户尤为重要。

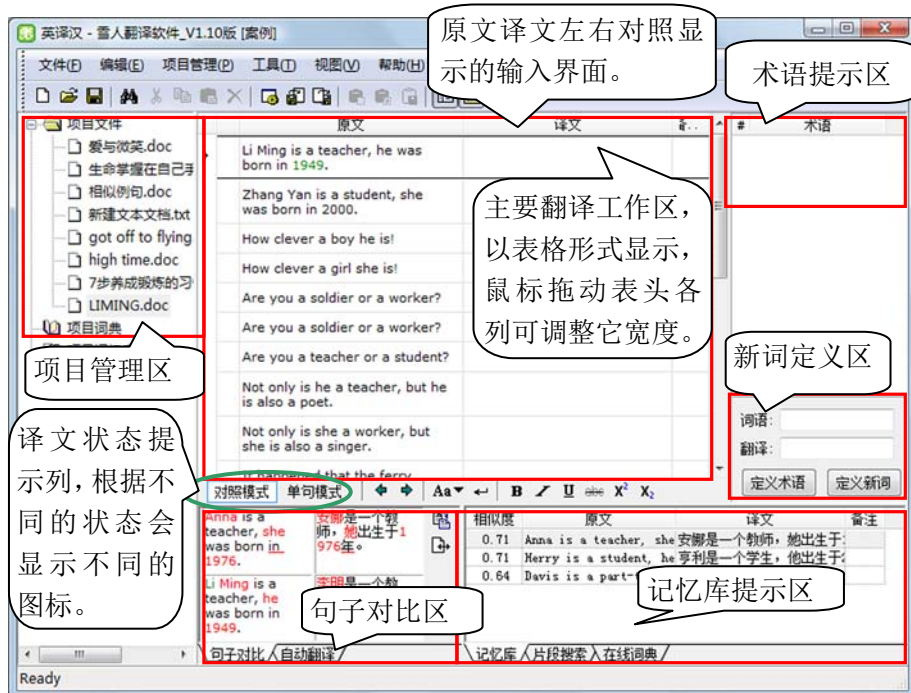
用户可以灵活运用翻译项目模板的功能，根据实际的工作情况决定是只使用一个模板，还是根据不同分类建立多个模板。

## 二、在对照模式中翻译

为满足不同用户的输入喜好，雪人软件支持左右表格排列的“**对照模式**”和“**单句模式**”两种输入模式，用户可以根据个人习惯采用，两种输入模式可以随时切换。导入需要

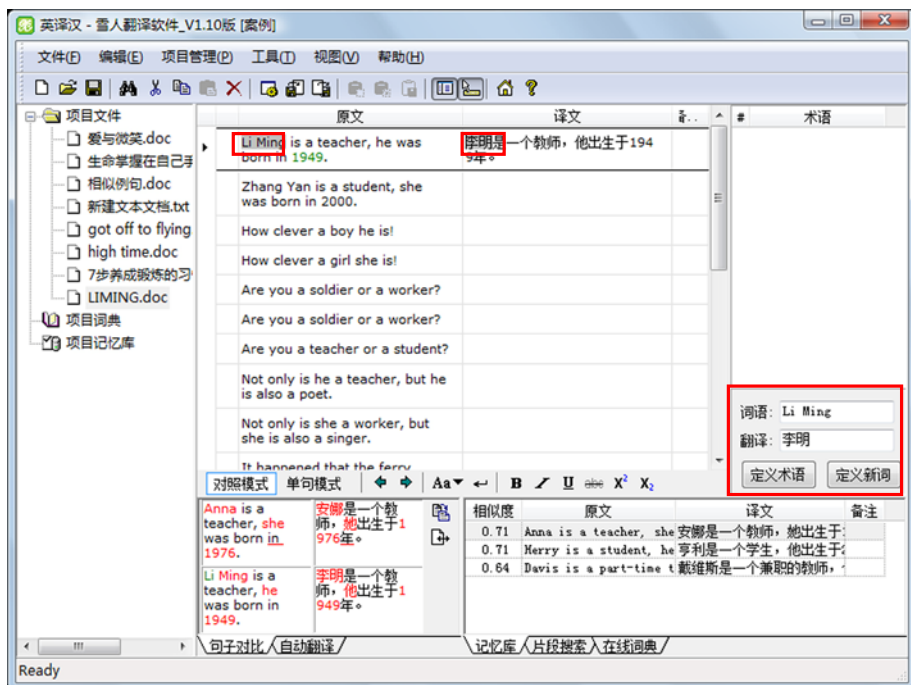
翻译的文件后，默认的界面为左右表格排列的“对照模式”，如下图：屏幕左侧是“项目管理区”。接着是“状态提示列”，它会根据不同的状态会显示不同的图标，以便他人校稿、修改。中间是“翻译工作区”，以表格的形式显示，有译文状态提示列、原文列、译文列、备注列，用鼠标拖动表头各列可调整它的宽度。右侧是“术语提示区”，遇到是术语的词语将会在此显示，同时在原文会以蓝色字体标注。右下方是“新词定义区”，用于在翻译过程中随时增添新词或术语。屏幕右下方是“记忆库提示区”，有“**记忆库**”、“**片断搜索**”和“**在线词典**”三个页签，“**记忆库**”页签显示的是从记忆库中搜索到的与当前句子相似的例句，并根据相似程度依次排列，最相近的排在最上面；“**片断搜索**”页签用于显示从记忆库搜索到的某个片段、短语，当鼠标在原文中划选了某几个词语后才会出现搜索的结果；“**在线词典**”页签当译员在原文中双击某词语或用鼠标划选词后，立即显示查找的结果，而且可以非常方便地在各个在线词典之间切换。屏幕左下方是“**句子对比**”和“**自动翻译**”，“**句子对比**”页签对于当前句，雪人 CAT 软件会从记忆库中自动找出参考例句与当前句进行对比；“**自动翻译**”页签显示的是当前句子的机器自动翻译的译文，译员可在此基础上修改或直接点击使用。


“对照模式”输入界面：原文、译文以左右对照的表格方式排列，这种界面简洁高效，条理清晰，对表格和审稿、校稿尤其方便。




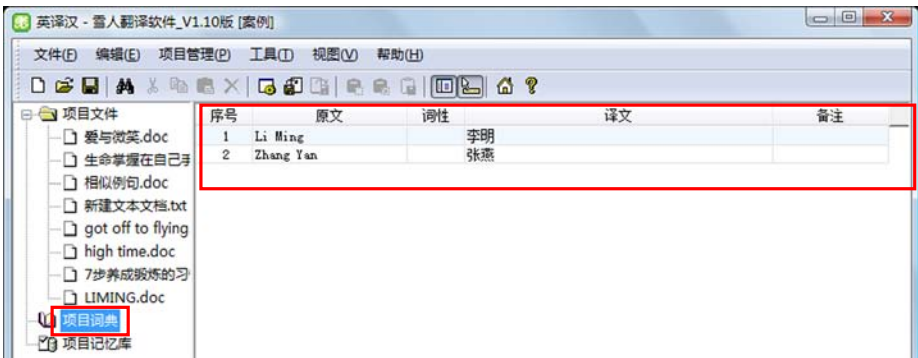



下面我们翻译第一句“Li Ming is a teacher, he was born in 1949. 李明是个教师，他生于1949年。”并利用鼠标在原文区选择了“Li Ming”，在译文区同时选择了“李明”，这是我们会发现这两个词语会自动出现在右下角的“新词定义区”，如下图：点击定义新词按钮或按【Ctrl+W】可以将“Li Ming 李明”作为新词添加到项目词典中。翻译好该句后，按回车键跳到下一句，此时会自动将刚才翻译好的句子保存到“项目记忆库”中。

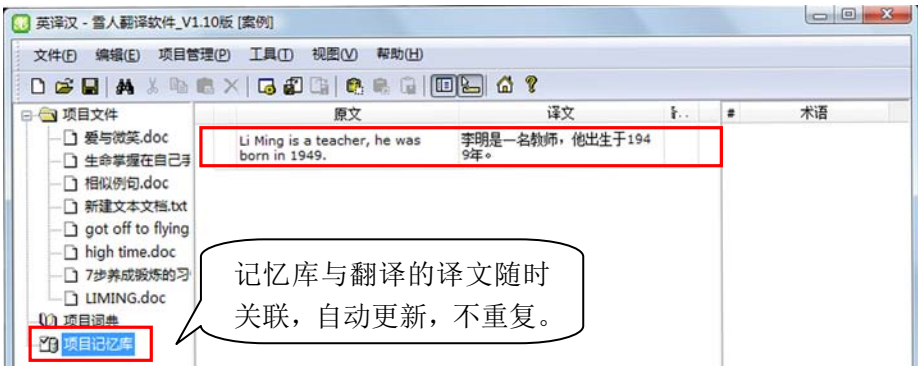


下一句“Zhang Yan is a student, she was born in 2000.”与刚才那句的句型一样，但有几个单词不同。由于“Zhang Yan”是个人名，在我们的词典库中没有这个词，我们可以在右下角的“新词定义区”直接通过输入来定义“Zhang Yan 张燕”，即在“词语”输入框中输入“Zhang Yan”，在“翻译”输入框中输入“张燕”，然后点击定义新词按钮。这时，点击界面左侧的 **项目词典**，我们看到刚才添加的两个新词，如下图：

 **项目词典**用于保存在翻译过程产生的词汇，一开始为空，用户可以点击右键导入或在翻译过程中积累；例如对于客户给定的术语，可以导入到项目词典后，利用【**右键菜单→设为术语**】功能将它们设定为术语（设为术语后以黑体显示，术语会自动出现在术语提示区，要求整片文章使用统一的译文）。

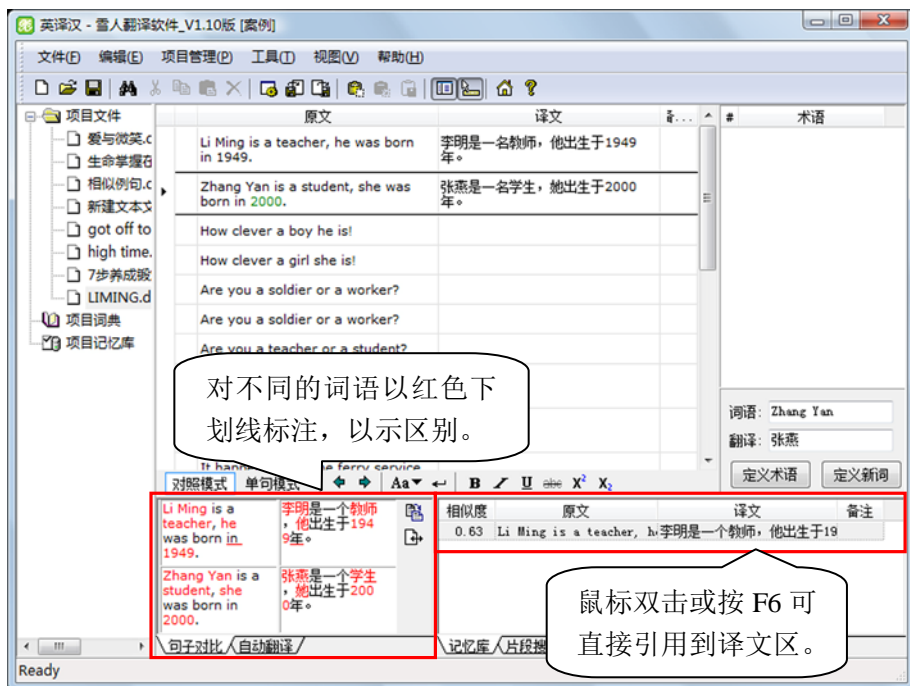


点击界面左侧的 **项目记忆库**，可以看到刚才翻译好的那一句已经自动保存到记忆库中了，如下图：雪人项目记忆库出色之处有二：其一、自动生成；其二、随时与原译文保持关联，自动更新，不重复。例如检查时发现某句翻译有错并进行了修改，记忆库中的那句会自动更新；而一般的 CAT 软件会保存了 2 句，然后要求译员再去记忆库中删除错误的那句。



现在让我们回到翻译窗口继续翻译 “Zhang Yan is a student, she was born in 2000.” 这一句，如下图：在屏幕下方的“记忆库提示区”软件自动从记忆库中找到了与当前句子相类似的例句，**用鼠标双击或按【F6】键可直接引用到译文区**，在双击下来的瞬间，我们可以发现软件已经对不同的词语进行了自动转换，生成了准确的译文“张燕是个学生，她生于 2000 年。”而且对不同的地方，在句子对比区中以**红色字体**显示，提示译员注意。这就是雪人翻译软件先进的 EBMT 技术的一个体现，这样可以使记忆库的例句价值得到更好的应用；类似上述的句子在一般的 CAT 软件中会被处理成两个不相关的句子，亦不可能根据记忆库的例句自动替换相应的词语生产准确的译文。

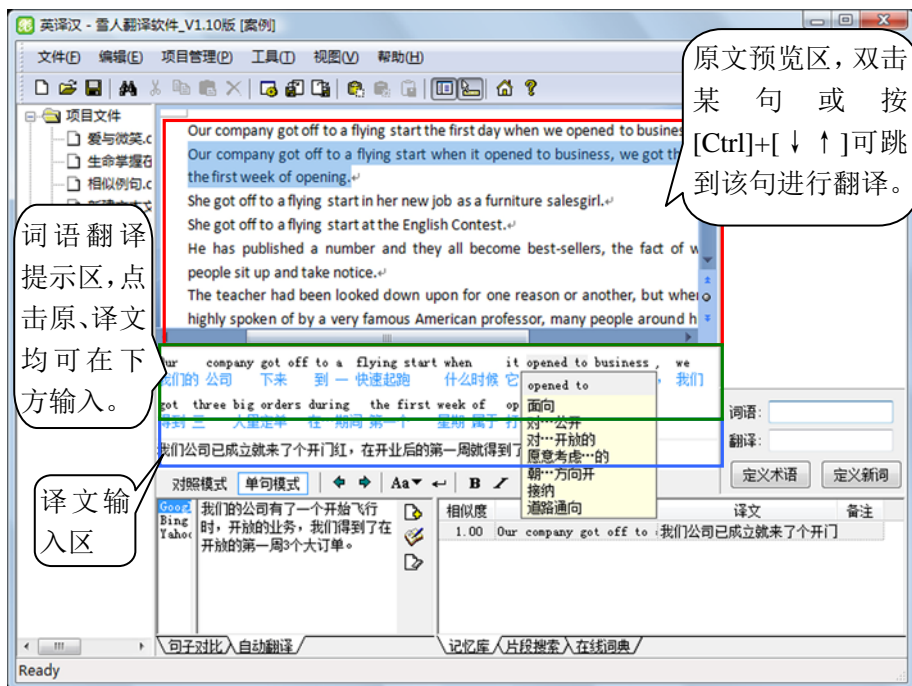




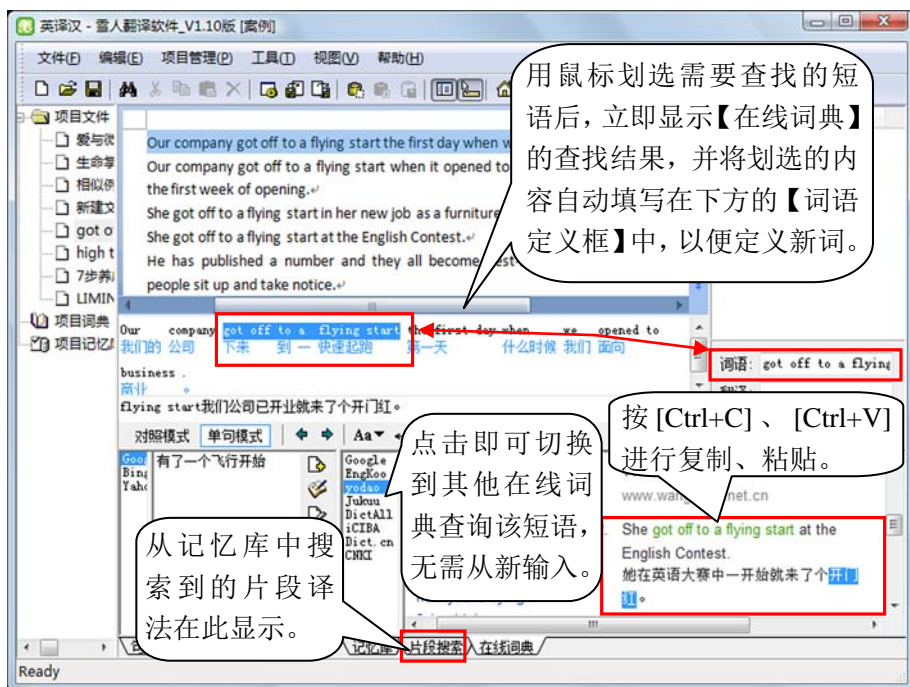
同理，如果已经翻译好了：“How clever a boy he is! 他是个多么聪明的男孩呀！”在翻译下一句：“How clever a girl she is!”只需在屏幕下方双击提示例句软件就可自动将它翻译为：“她是个多么聪明的女孩呀！”。

### 三、在单句模式中翻译

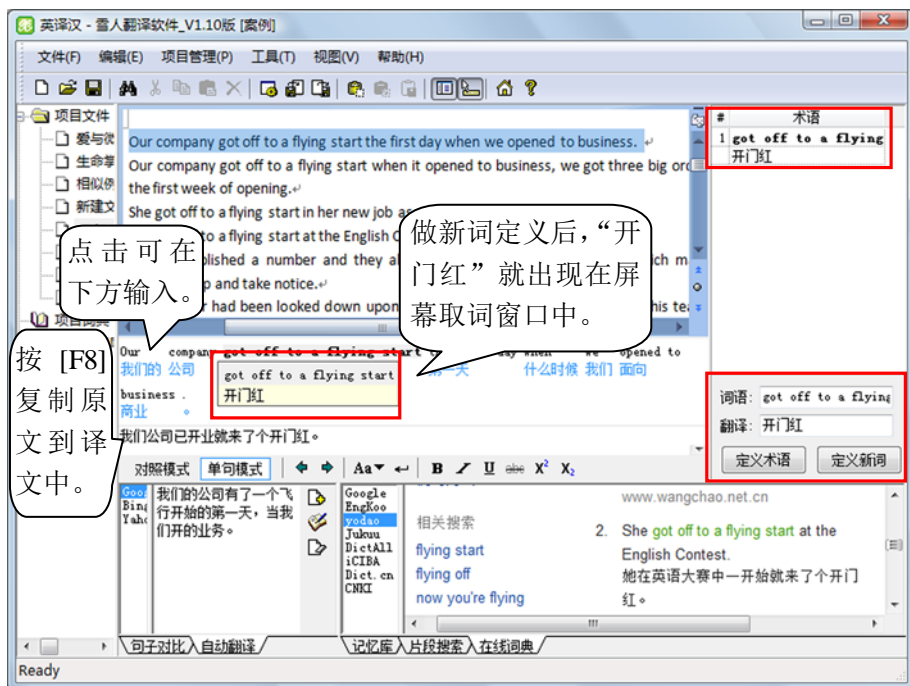
刚才讲述了在**对照模式**中如何进行翻译，下面将讲述在**单句模式**中如何翻译？两种模式各有长处，而且可以随时切换，切换后光标会自动定位在当前翻译的句子中。点击“**单句模式**”按钮，界面如下图：在屏幕的中部是“原文预览区”，在此可以图文并茂地预览原文，其下方是“词语翻译提示区”，软件自动将当前句的各词翻译列出，点击它即可在下方输入，如果默认的取词不准确，可以将鼠标指向该词，这时会出现一个下拉屏幕取词窗口，列出该词的其他译法，点击相应的解释即可在下方输入。



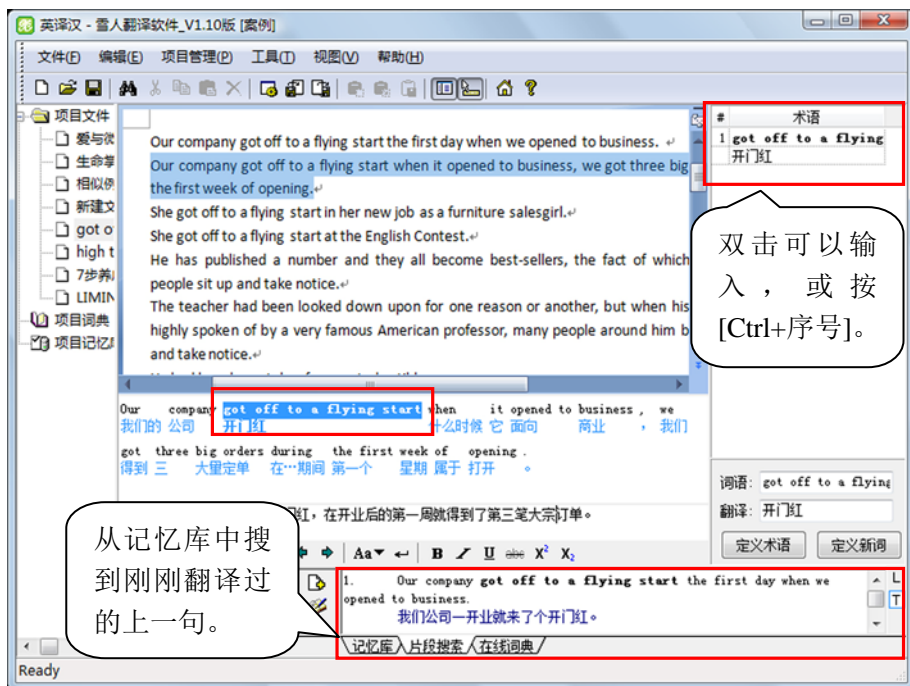
下面我们来翻译：“Our company got off to a flying start the first day when we opened to business.” 这一句，如上图，在“词语翻译提示区”中我们可以看到自动翻译对“got off to a flying start”短语的翻译不正确，该短语如何译呢？我们有两种海量资源可以利用，一种是从用户自己积累的记忆库中去搜索该短语的曾经译法，即利用软件强大的“片段搜索”功能；一种是利用位于屏幕右下方的“在线词典”功能。



当我们用鼠标划选了“got off to a flying start”短语后，如果用户的记忆库中曾经翻译过这个短语，此时在屏幕下方的“片段搜索”区会出现从记忆库中搜索的结果；这里我们没能从记忆库中找到该短语的翻译，用户可选用“在线词典”功能，单击在线词典按钮，屏幕的右下方立即出现该词语的译法，我们选择“在线词典”中的“开门红”这几个字后，点击【Ctrl+C】进行复制（如上图），接着在屏幕右方的“词语定义”区的“翻译”输入框中点击【Ctrl+V】进行黏贴，将它粘贴过来，然后点击定义术语按钮或定义新词按钮将该短语的翻译保存到“项目词典”中，在翻译过程中若遇到生词，不管是利用“在线词典”还是“片段搜索”都可以非常方便地随时增加新词或术语。上述当我们将“got off to a flying start 开门红”做了术语定义后，这时在“词语翻译提示”区中该短语的翻译立即更新为“开门红”（如下图），而且该短语会立即出现在屏幕取词窗口。



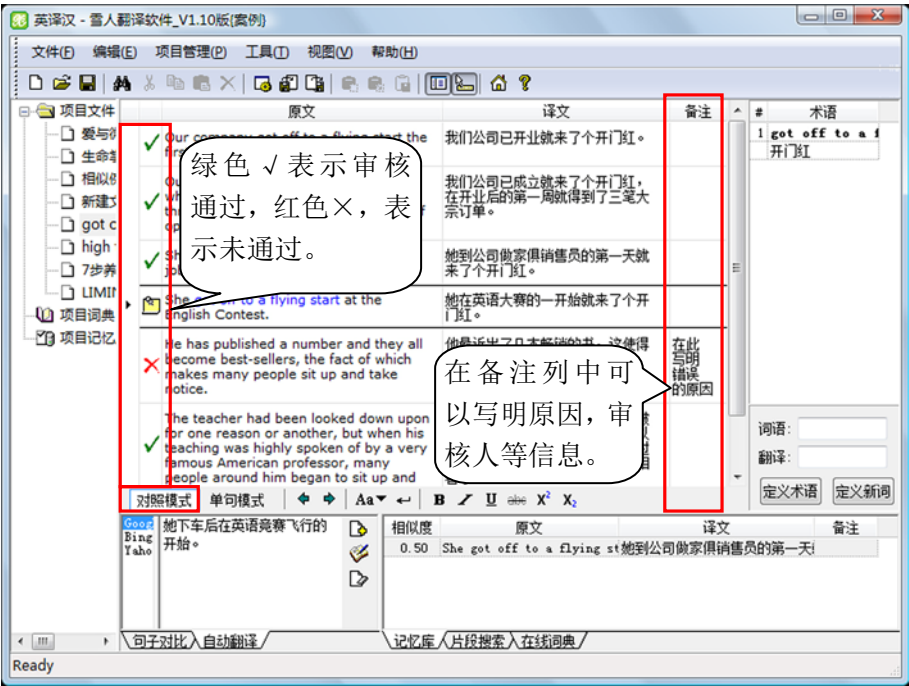
翻译好上一句后，按【Ctrl】+【↓】键翻译下一句或用鼠标双击（【Ctrl】+【↑】翻译上一句），例如：翻译“Our company got off to a flying start when it opened to business, we got three big orders during the first week of opening.”由于刚才我们将“got off to a flying start 开门红”做了术语定义，所以，该句的这个短语已经自动翻译为“开门红”了，而且，刚才翻译的“Our company got off to a flying start the first day when we opened to business. 我们公司一开业就来了个开门红。”这一句雪人 CAT 软件已经自动保存到用户记忆库中，此时当鼠标标选了“got off to a flying start”这个短语后，在屏幕下方的“片段搜索”窗口中，已经将刚才翻译的这一句自动搜索出来了；同时在右侧的术语区也出现了该术语的提示，双击可以输入，或按【Ctrl】+【1~9 序号】直接复制术语到译文中。如下图：



对照模式和单句模式各有各的优势，而且切换非常方便。对于句子的拆分与合并，则必须在对照模式中进行，单句模式中不支持。在对照模式中，把光标移到需要拆分的地方，按【Enter】键，即可把当前的句子拆分为两个单元句子；把光标移到需要合并的第一句话的末尾，按【Delete】键就可合并两个单元句子。

## 四、审稿、校稿

对照模式在审稿、校稿方面尤显方便。若初次翻译工作完成后，可以将“项目文件”发给审稿人进行审核，雪人翻译软件将涉及当前项目的所有内容打包在一个“项目文件”里，其中包括该在翻译过程中形成的项目词典和项目记忆库等等，译员只需要将“项目文件”复制给他人即可。审核人打开他人发的项目文件后，如果觉得译文准确，就在当前行按下【F9】键，软件会在此行的第一列标注一个绿色的“√”，表示审核通过；如果觉得某句有问题，就在当前行按下【F4】键，软件会在此行的第一列标注一个红色的“×”，表示未通过，并且可以在备注列中写明未通过的具体原因。审查完成后将该“项目文件”发回给原译员。





原译员打开审核人发来的“项目文件”，看第一列的图标及备注信息，就知道存在哪些问题，对某句修改好后，按下【F4】键原来的×会变成黄色图标，表示修改完成。按【Ctrl+N】可直接跳到下一个红色“×”位置继续修改，【Ctrl+N】对修改人员来说极为方便，不必费神去查找还有哪里需要修改。

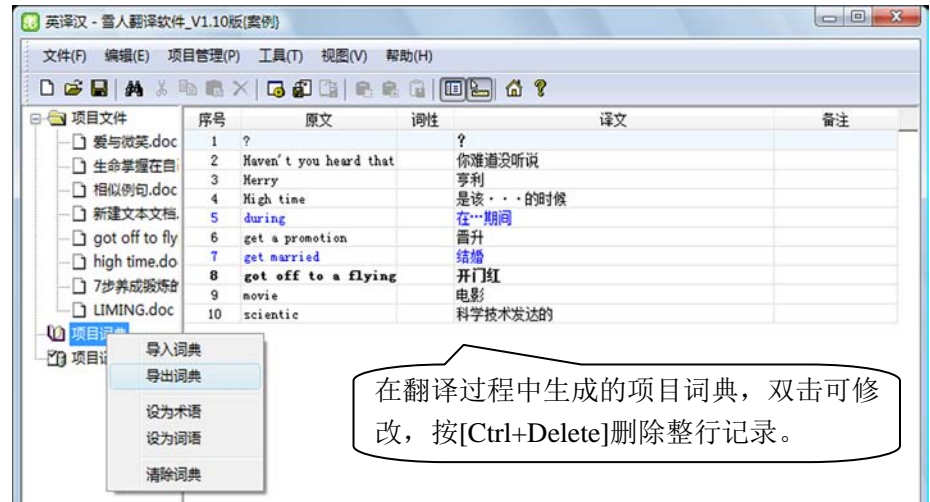
**对照模式**中的表格各列的宽度可以随意调整，调整方法：将鼠标移向表头处拖动即可改变宽度。若屏幕空间不够，可以向右拖动备注列将它隐藏，待需要时再向左将它拖出来。



五、项目词典和记忆库的维护

翻译工作完成后，在翻译过程中生成的词典及记忆库会自动保存在项目词典和项目记忆库中，用户可以进行再次整理、编辑，编辑好后可以导出为外部的词典或记忆库，日后在其他翻译项目中可以通过“项目设置”的“添加”来引用，而不是将它导入到新项目的  项目词典 或  项目记忆库。因为“添加引用”的方式可以更好地减少项目文件的大小和对系统资源的占用，对配置不高的机器尤为重要。

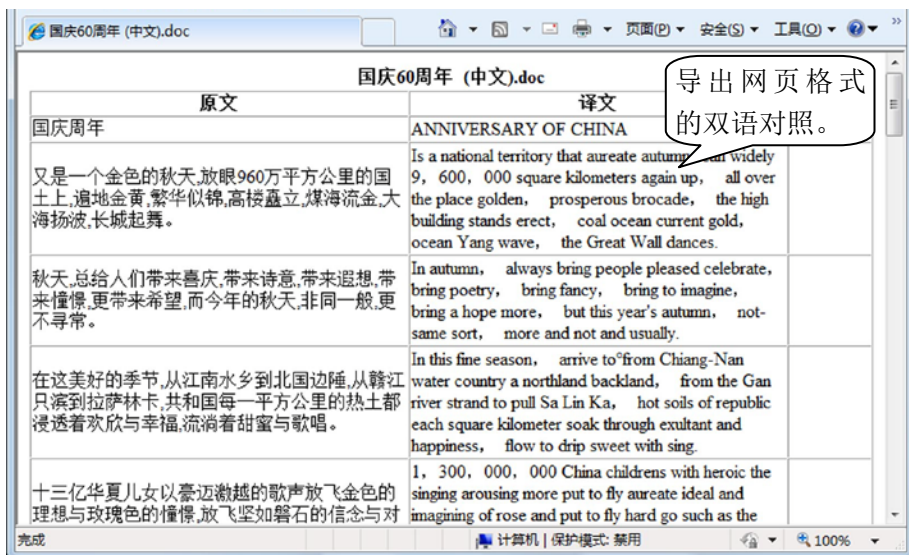
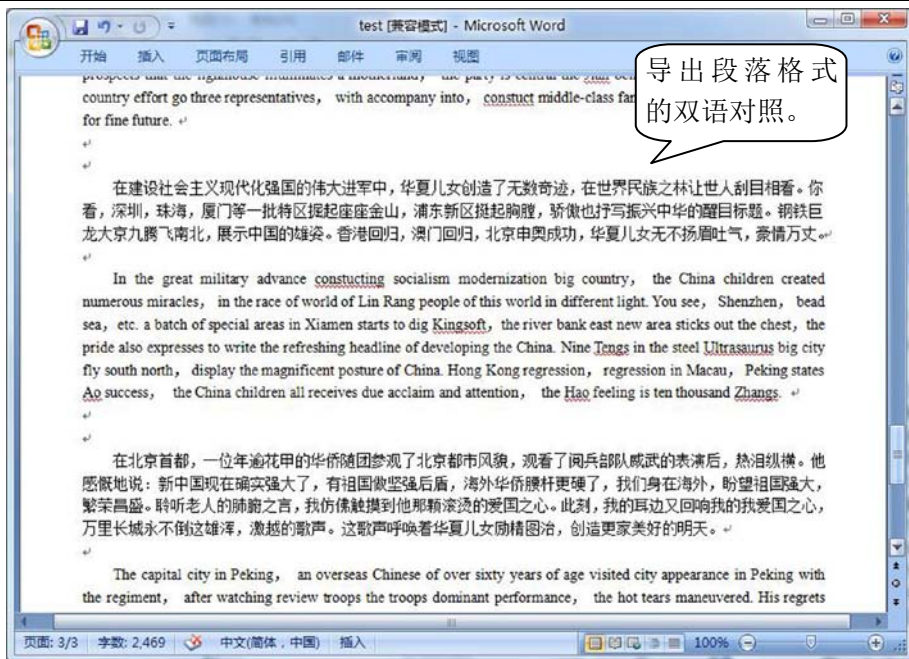
软件默认状态是长光标的浏览状态，在浏览状态不能修改，按【F2】可以进行编辑和浏览状态的切换，用户既可以对译文进行修改，亦可以对原文进行修改。在浏览状态下按【Ctrl+Delete】键可以删除整行记录。修改完成后，点击右键菜单的导出功能可以将它导出为雪人格式的外部词典或外部记忆库，以后的翻译项目就可以直接引用它了（标准版除导出为雪人格式的文件外，还可将记忆库导出 TMX 和 EXCEL 等多种格式，其中，TMX 格式是记忆库翻译软件记忆库交换的标准格式，免费版只能导出雪人格式的。如下图：



六、导出译文或双语对照

翻译、校对工作都完成后，点击左侧“项目文件”下的具体文件，利用【右击菜单→导出译文】功能将译文导出。雪人可以导出与原文严格一致的译文，而且还可以导出双语对照文件，双语对照有两种格式：段落格式和网页格式；段落格式即一段原文、一段译文，网页格式以句为单位，一句原文、一句译文。点击【右击菜单→导出双语对照文件】功能即可导出。另外，文件保存时记录翻译的位置，在打开文件时转到保存的位置。

段落格式的双语对照和网页格式的双语对照分别如下图：



综上所述,使用雪人翻译软件无论从最初的安装,到整个翻译工作完成,都是相当的简单、易用,而且软件功能强大、搜索速度快,即使百万级别的记忆库也可以轻松应对。




## 第三章 软件常用功能介绍


### 第一节 句子的拆分与合并

虽然一般情况下雪人软件在导入原文时可以对句子进行正确的拆分，但是我们仍然向译员提供了方便的句子合并和拆分功能。把光标移到需要拆分的地方，按【Enter】键，即可把当前的句子拆分为两个单元句子；把光标移到需要合并的第一句话的末尾，按【Delete】键就可合并两个单元句子。

**注意：**拆分或合并句子功能仅限于在**对照模式**中，**单句模式**中不支持，需要切换回**对照模式**下进行。译员可根据实际需要自由地进行句子的拆分与合并，方便进行翻译工作。

### 第二节 定义术语和定义新词

术语一般是指专业的词汇，有些翻译项目，对方会提供术语库，这时，将它导入到  **项目词典** 中，利用【**右击菜单→设为术语**】功能将它们设定为术语就行了。术语是在整个翻译项目中需要统一译法的词，一般情况下一个术语只对应于一种解释，定义为术语的词在翻译时会以**蓝色**提示，同时会自动出现在右侧的术语提示区。定义新词指的是定义为普通的词汇。

目前大多数译员是一个 CAT 软件加一个词典软件协同工作，而且还存在有些 CAT 软件和词典软件冲突的问题，对查到的词语只能再次录入或复制。雪人除了系统自带巨量的词典外（标准版有 60 多个专业的、1200 万的词典，免费版有 30 万词），还有一个突出优点就是定义新词非常方便，若配合“在线词典”和“片段搜索”功能一起使用，则更显高效！即：用鼠标划选生词后，“在线词典”和“片段搜索”窗口即可搜索出结果，点击屏幕右下角的  **定义新词** 按钮即可。对定义了词或术语，当鼠标指向某词或短语停留片刻，立即会弹出屏幕取词窗口，单击某项合适的译文即可输入；而且在单句模式中会以此词义作自动翻译。这些是用其他的词典软件协同工作无法达到的效果！



#### 定义新词的方法有两种：

方法一：


1. 用鼠标在原文处双击某个词语或划选一个片段，单句模式下需要在下方的“词语自动翻译区”中划选；
2. 在译文中划选其相应的意思或直接在右下角的“词语定义区”中输入；


3. 按“添加新词”按钮或【Ctrl+W】即可将选定的词保存到定义新词中；

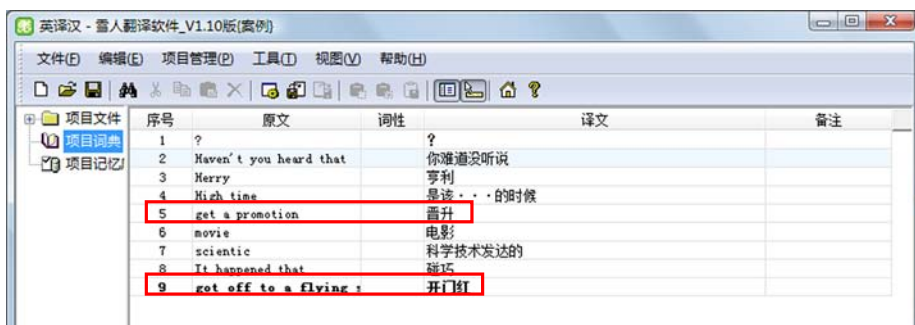
方法二：

1. 用鼠标在原文处双击某个词语或划选一个片段，单句模式下需要在下方的“词语自动翻译区”中划选；
2. 在“自动翻译窗口”立即出现其译文，若不确切，还可以对译文进行修改；
3. 点击“自动翻译窗口”旁边的定义新词可将它加入到项目词典。

## 定义术语的方法有三种：

方法一、点击屏幕左侧的项目词典，选择想要定义为术语的词汇，按下【F5】键即可定义完成。

方法二、通过鼠标右击项目词典，选择【右键菜单→设为术语】可以将此项目词典的全部词汇设定为术语，设置成功后，整个词库以粗黑体显示，如下图（注：项目词典包括术语和新词，术语以粗体显示、新词不加粗，以示区分。）这个功能对导入客户的术语库尤显方便，即首先将客户的术语库导入到项目词典中，然后利用【右键菜单→设为术语】一次过全部将它们设为术语即可。



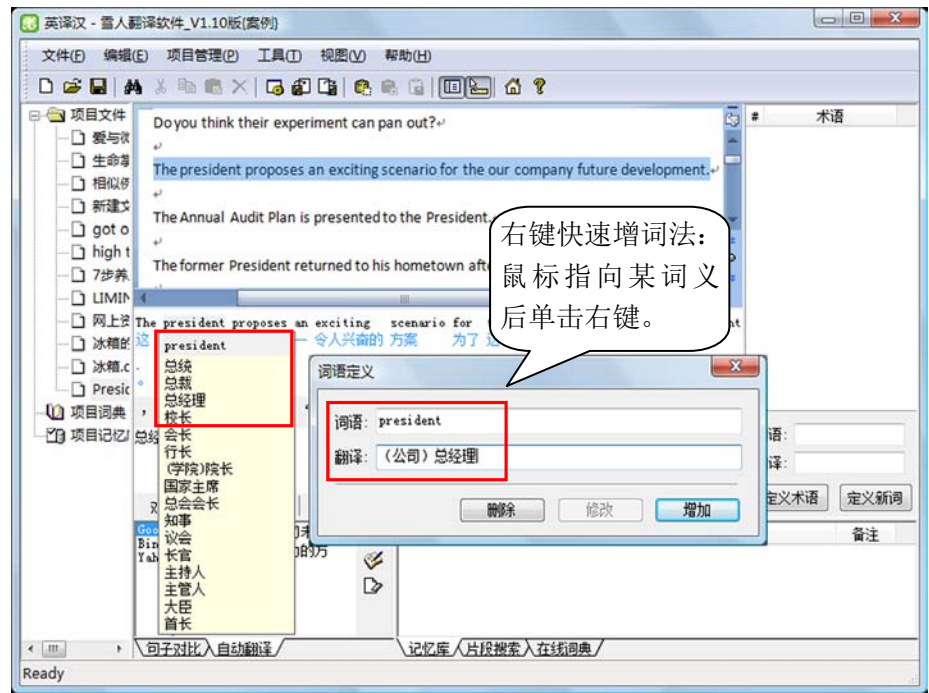
方法三、在翻译过程中，译员可以随手收集词语，只要用鼠标在原文区选择需要收集的原文和在译文区选择对应的译文，它们就会分别出现在右下角的新词定义区中，点击定义术语按钮或按【Ctrl+D】键即可添加术语。添加的词条在其它的词条的末尾。词库里的词语信息是可调整的，也可以删除，译员可通过单击该词条，按【Ctrl+Delete】键删除该记录。

**建议：**若在翻译时能多对一些短语、片语做新词定义，这样日积月累下来，不但可以加快输入的速度，更可以很好地提升自动取词的准确度。

### 第三节 自动调整取词顺序

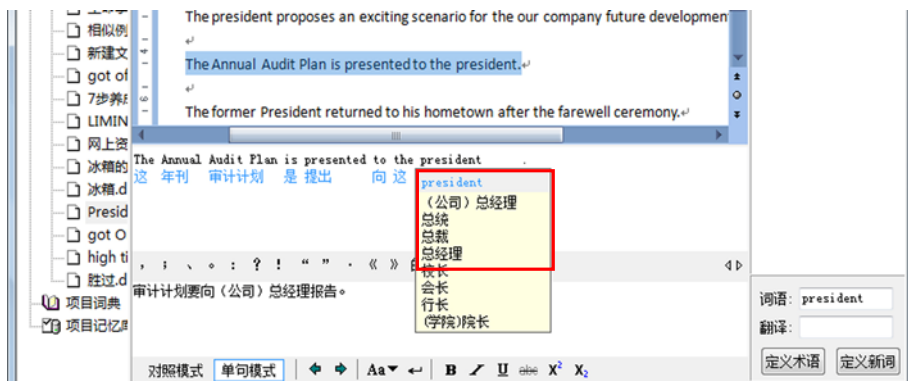
由于存在一词多义，软件在自动翻译时默认取词典的第一个意思，这样难免会出现取词不正确的情况。但是只要某词义一旦被点击使用过，则软件立即会将该词义排在屏幕取词窗口的第一位置上，本翻译项目中若再次遇到该词时亦会自动按此词义自动翻译。


例如下图中：“President”在词典中的第一意思是“总统”，由于默认取词典的第一个意思，所以，软件在词语自动翻译时将文章中所有的“President”都翻译为“总统”；但在我们这篇文章中“President”应该翻译为第三个意思，即“总经理”，翻译时我们可以点击屏幕取词的第三总经理”；如果觉得该词义还是不够准确，可以以此作参考增加新词，称之为“右键快速增词法”：将鼠标指向屏幕取词窗口的“总经理”后（此时该词会显示为蓝色），单击右键，会弹出“词语定义”窗口，并自动将原文、译文填写好，用户在此基础上修改即可，例如下图中：增加“（公司）”二字，改成“（公司）总经理”。



然后按**增加**，即可将该词的原义“总经理”修改为“（公司）总经理”（在用户词典中亦增加了该词义）。接着按**【Ctrl+↓】**键翻译下一句，当再次碰到“President”这个词时，软件已经自动将它翻译为“（公司）总经理”了，而且在弹出的屏幕取词窗口中，“（公

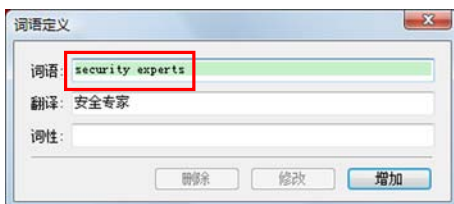
司) 总经理”的解释排在了第一位,“总统”的解释则列在第二位了(如下图: )。



另外,在右键弹出的“词语定义”窗口中,还可以直接对“用户词典”里的词进行删除或修改,但是对于外部通过添加引用的词典,删除、修改按钮为灰色,表示不可修改。如果有时不希望自动弹出屏幕取词窗口,可以按下工具栏的  鼠标取词按钮。

## 大写转小写快捷键【Ctrl+L】


鼠标划选词语后单击右键,会弹出“词语定义”窗口,如下图:光标在输入框中,按【Ctrl+L】可以将大写字母转换为小写。

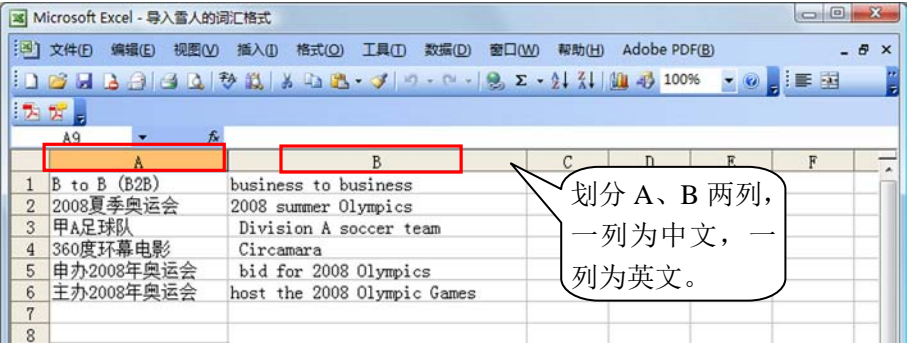


## 第四节 自己建词库(项目词典)

除了使用雪人系统自带丰富的词典外，用户还可以创建自己的词库。当新建一个翻译项目后会自动生产一个空的“项目词典”，用于存放用户在翻译过程中添加的词语和术语；此外，用户也可以在“项目词典”导入其他格式的词典文件，创建自己的词库。

### 一、项目词典导入

右击屏幕左侧的 **项目词典** 会弹出一个菜单，选择【右击菜单→导入词典】功能可以导入词典，雪人翻译软件导入词典的速度非常快，而且**会自动删除重复的词汇**。雪人支持导入词典的格式：**\*.xls、\*.txt、\*.cxv、\*.xml、\*.xlsx、\*.yxd (雅信格式) 和 Dic (雪人格式) 的词典**；若是导入 EXCEL 格式的词典，则只需在 EXCEL 中分 A、B 两列，一列为中文、另一列为英文，就可以将 EXCEL 格式的词典导入到雪人中了。如下图：



若是文本格式的词典，则在中英文之间需要设置分隔符，分隔符建议采用某些特殊的，例如：Tab 制表符、|| 等等。设置好分割符后，就可以导入了。如下表：

B to B (B2B)		business to business
2008 夏季奥运会		2008 summer Olympics
甲 A 足球队		Division A soccer team
360 度环幕电影		Circamara
申办 2008 年奥运会		bid for 2008 Olympics
主办 2008 年奥运会		host the 2008 Olympic Games

点击【右击菜单→导入词典】功能，导入到雪人中，如下图：雪人给每个词条排上序号，方便查看。在此导入的词典仅供当前翻译项目使用。单击标题栏可以调整词典的排列

方式，用户可以根据需要将词典按原文、译文或备注的顺序或倒序排列（在同一标题栏点击可以设置顺序或倒序的方式）。



## 二、项目词典的导出

成功导入后，可以进行编辑，然后将“项目词典”中词汇导出雪人 DIC 词典格式；或者完成翻译工作后，在“项目词典”中对本次翻译积累的词汇进行编辑、整理词库里的词条。软件默认状态是长光标的浏览状态，在浏览状态不能修改，按【F2】可以进行编辑和浏览状态的切换，直接双击亦可以进行修改，译文和原文都可以修改。在浏览状态下按【Ctrl+Delete】键可以删除整行记录，其他更多的词典维护功能可以点击右键菜单。整理完成后，右击屏幕左侧的 **项目词典**，选择【右击菜单→导出词典】功能可以将词典导出；词典导出后即作为外部词典，以后任何翻译项目都可以根据需要调用它。

## 三、在新的项目中引用

译员在以后的翻译中根据需要添加不同的词典进来，可点击【项目管理→项目设置】功能，然后点击 **用户词典** 标签页，再单击 **添加** 按钮，就可以在新的项目中引用词典。如下图：用户可以根据专业将词典分类，不同专业的词典分别导出并通过文件名区分。这样在引用时就非常方便了，选用哪个专业的词典，就添加哪个专业的词典进来，雪人可以支持添加任意多个的词典。

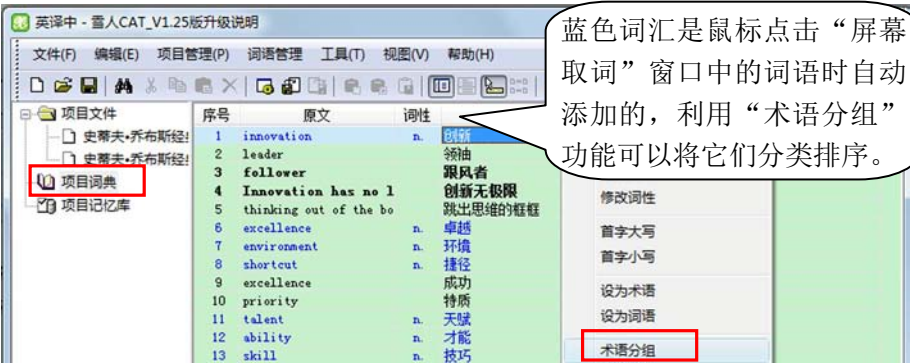




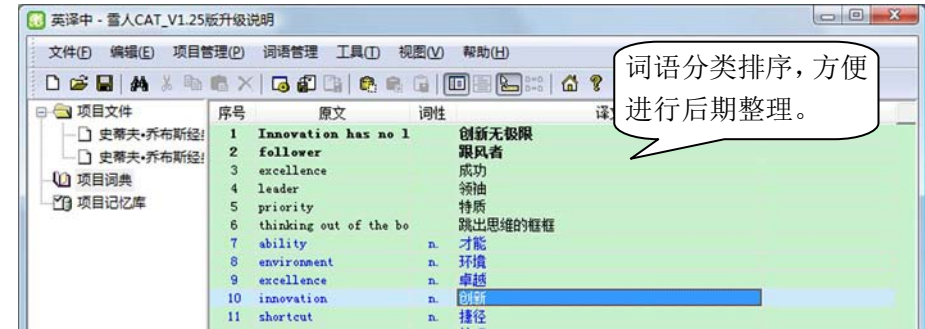
注意，词典是单项的，例如，需要同时建立汉英词典和英汉词典，该如何实现？则具体的做法是：在雪人中新建一个汉译英的翻译项目，在屏幕左侧的“项目词典”中点击右键菜单，导入汉译英的词汇，然后将词典导出为 DIC 格式，则这个汉英的词典就做好了。然后，再新建一个英译汉的翻译项目，在“项目词典”中导入刚才导出的那个 DIC 格式的词典文件，在导入的同时，雪人可以自动做词典反向处理。

四、项目词典中词汇整理

用户翻译时，利用“屏幕取词”窗口点击输入的词语，软件会自动添加到“项目词典”中，并以蓝色字体显示。翻译完成后，通常会对本次翻译积累的词汇进行整理，例如导出等，而对软件自动添加的蓝色词汇若不需要导出时，由于默认状态下，蓝色词汇是与其他词汇混排在一起的，这样就给整理工作带来了麻烦，如下图：



利用【右键菜单→术语分组】功能，可以将术语（黑色粗体）、一般词汇（黑色）、自动添加词汇（蓝色）区分开来，术语排在最前面，接着是一般词汇，最后是软件自动添加的（蓝色）词汇。如下图：此时如果不需要这些（蓝色）的词汇，可以按住【Shift】键点选这些词语，然后删除。注意：删除后，可能会改变软件在当前文档的“屏幕取词”窗口中的词义排列顺序。



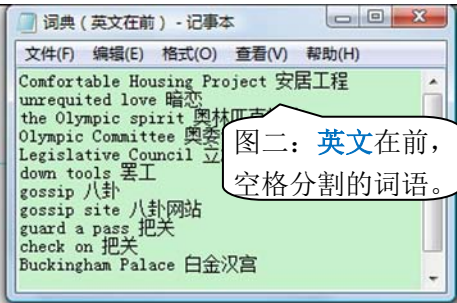
## 五、整批修改词性

在项目词典中，选择多个记录并点击【右键菜单→修改词性】功能，可以整批设置所选的词语的词性，如下图。



## 六、反向匹配

用户在导入 TXT 格式的词典或者记忆库时，通常会遇到下图所示的 TXT 原文件（这里以词典为例说明，记忆库同理）：（图一）的中英文间是以空格分割的，且中文在前，英文在后面；这种格式雪人在导入时会自动正确区分原文、译文。

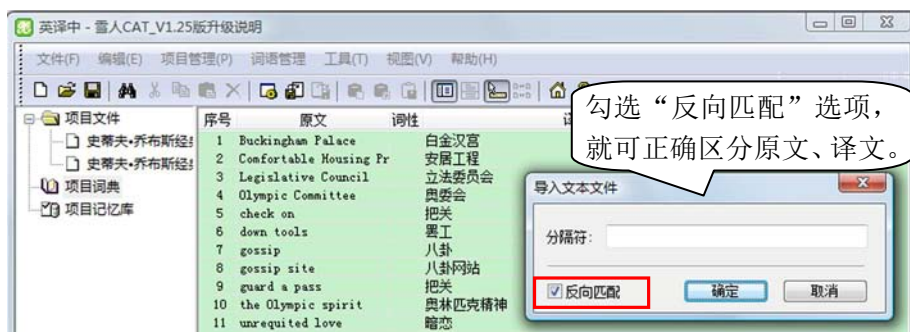




而（图二）也是以空格分割的词语，不同的是**英文**在前，中文在后面；由于英文的单词之间也是以空格分割的，在以前的版本中，这种格式的词典就不能正确导入了（如下图）：



以前用户需要手动或使用其他的分隔符工具，在 TXT 原文件的英文、中文之间插入分隔符，方能正确导入。现在增加了“反向匹配”功能，则可以很好地解决上述问题，软件会自动识别是否需要进行“反向处理”，如果需要，则会自动勾选“反向匹配”选项。即使是空格分隔的、英文在前的词汇或者双语句对，都能很好地自动区分中英文。如下图：



## 第五节 自己建记忆库（项目记忆库）

用户翻译过的句子，会自动保存在“项目记忆库”中形成记忆库。或者利用雪人的“双语对齐”功能，将收集到的双语资料进行句级对齐，制作成记忆库，关于“双语对齐”的具体用法请参看“第十章 双语对齐创建大型记忆库”。还可以在“项目记忆库”中直接导入其他格式的文件，制作成记忆库，雪人目前支持导入记忆库的格式有：**TMX(CAT 软件记忆库交流标准格式)、EXCEL、TXT、CSV 格式和 STM(雪人的记忆库格式)**；文本格式的导入非常灵活，既可以自动识别分隔符，也可以手工指定分隔符。在导入文本格式时，软件会自动识别分隔符，并将识别到的分隔符提示给用户，用户可以修改；如果原文中存在分隔符不统一的情况，则软件不能正确识别，这时需要用户在提示窗口中手工输入分隔符。

在“项目记忆库”中点击**【右键菜单→导出记忆库】**功能，可以将记忆库导出，记忆库文件是双向的，同一个文件可以用于汉英和英汉。记忆库支持导出的格式有：STM(雪人的记忆库格式)、EXCEL、TMX 格式，免费版的用户只能将记忆库导出 STM 格式。雪人导入的速度非常的快，一分钟导入 15 万 TMX 格式的句对到记忆库，并且在导入时已经删除了重复的句子和统计出实际例句数量，确保译员自建的记忆库中同样的句子不会出现重复。

标准版的用户利用导入、导出 STM 格式的记忆库可以达到记忆库合并的目的，而免费版的用户由于不能导入 STM 格式的记忆库，记忆库的合并相对会麻烦一些，这里讲述一下免费版用户合并记忆库的两个方法：

方法一：

1. 新建一个翻译项目
2. 导入 EXCEL 格式或 TMX 格式的记忆库。免费版用户可以将已经做好的项目文件导出“网页格式的双语对照”文件，然后用 EXCEL 打开，即可变成表格形式的，在 EXCEL 中进行合并。然后再将 EXCEL 文件导入到雪人的项目记忆库中，即可实现记忆库的合并。

以后，需要再增加记忆库，则打开这个翻译项目，然后将 EXCEL 格式或 TMX 格式记忆库导入，新导入的会贴在旧的后面，这样就可以积累一个大的记忆库了。即：用建立一个“翻译项目”的方式来达到合并记忆库的目的。


方法二：

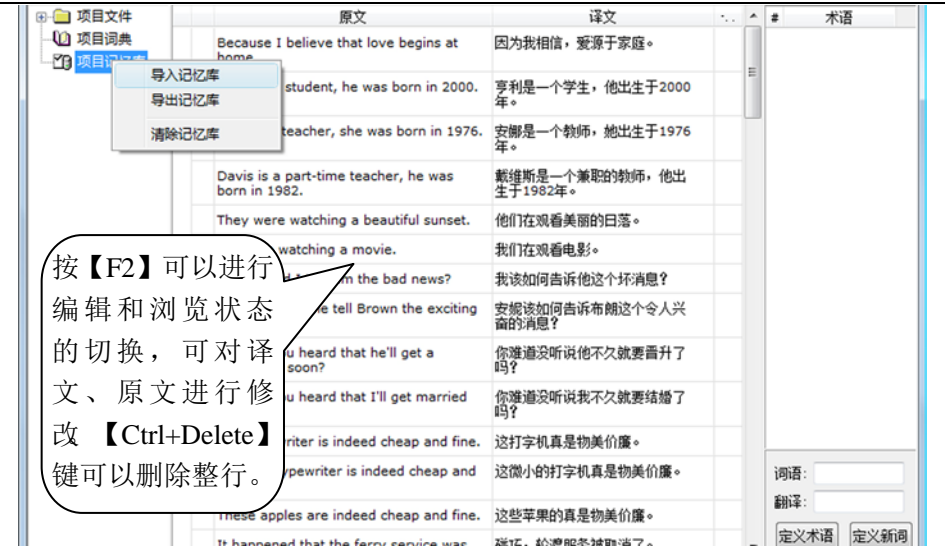
利用雪人在一个项目文件下可以导入 200 份文档的特性，就将所有要翻译的文档都导入到这一个项目文件来翻译，这样翻译好后形成的记忆库、词典就全在一起了。当超过 200 份时，可以将一些已经译过的文档删除掉，删除文档时注意选择保留记忆库选项，这样文档虽然删除了，但它产生的记忆库却永远留在了项目记忆库中！即：新建两个，一个是英译汉，一个是汉译英，以后将所有的英译汉项目就在这个“汉译英.stp”中导入来做（最

多可以导入 200 份，若满了可以将之前已经翻译过的文档删除)。同理，将所有的英译汉项目就在也在那个“英译汉.stp”中导入来做。实际上你只需要保存这两个项目文件即可。不过，若英译汉产生的记忆库希望在汉译英中使用，则需要将记忆库导出，然后在汉译英中通过“添加”记忆库来引用它，不需要导入。

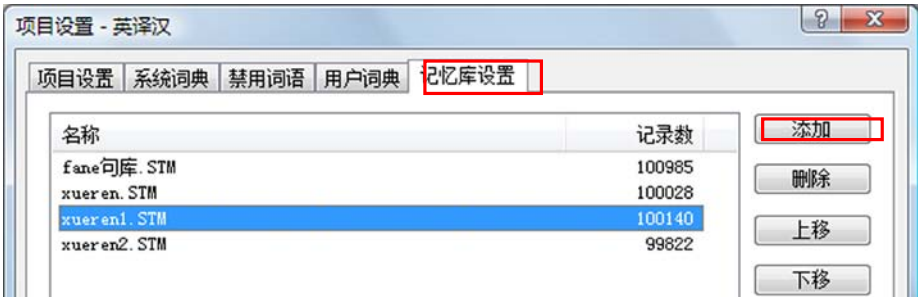
## 一、项目记忆库的维护

当新建一个翻译项目后亦会自动生产一个空的项目记忆库，用于存放用户在翻译过程中形成的双语句对。一般的 CAT 软件当用户翻译好某句后，则需要点击保存到记忆库按钮才会保存到记忆库中，保存后如果检查时发现翻译有误，又对译文做了修改，则又需要点击按钮再保存一次，而且记忆库中此时已经保存了两个记录了，对于先前不正确的记录，则需要用户自己进入记忆库中将它删除。**雪人记忆库的优点是：自动记录和自动更新，绝不重复！**雪人自动将用户翻译好的句子保存到记忆库中，而且自动与译文同步更新，即：若译员对原来的译文做了修改，记忆库的相应句子亦会自动更新，不会将修改后的句子当作新的例子重复记录，保证记忆库永不重复。记忆库由软件自动生成，用户可以再次编辑，软件默认状态是长光标的浏览状态，在浏览状态不能修改，按【F2】可以进行编辑和浏览状态的切换，**用户即可以对译文进行修改，亦可以对原文进行修改**。在浏览状态下按【Ctrl+Delete】键可以删除整行记录。

除了在翻译过程中自动形成记忆库外，也可通过导入外部数据库。导入记忆库的格式：**STM(雪人格式)、TMX、XLS、TXT、CSV** 等格式；导入速度极快，即使是数十万的句对，也可在瞬间完成，这给译员自己创建记忆库带来了极大的方便。导入的方法类似项目词典的导入，右击屏幕左侧的 **项目记忆库**，选择【右击菜单→导入记忆库】功能即可。记忆库里的原文、译文都可以修改，在记忆库中对句子的调整不会影响翻译项目里的源文和译文。对于不需要的句子，在浏览状态下可通过【Ctrl+Delete】键来删除（按【F2】键可以进行浏览和编辑状态间的切换）。



导出的记忆库格式为 STM 格式的文件，下次需要用上时，可点击【项目管理→项目设置】菜单，然后点击记忆库设置标签页，再单击添加按钮即可在新的项目中引用它了。

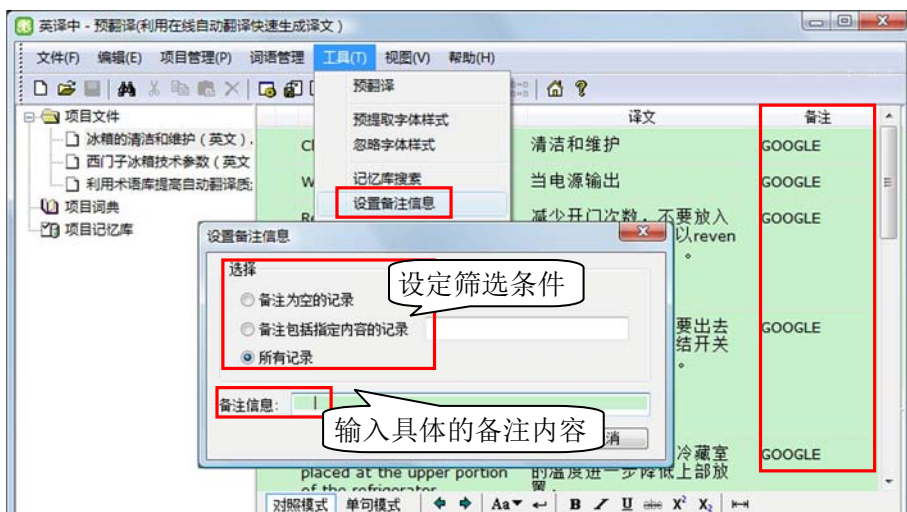


## 二、 整批设置备注信息

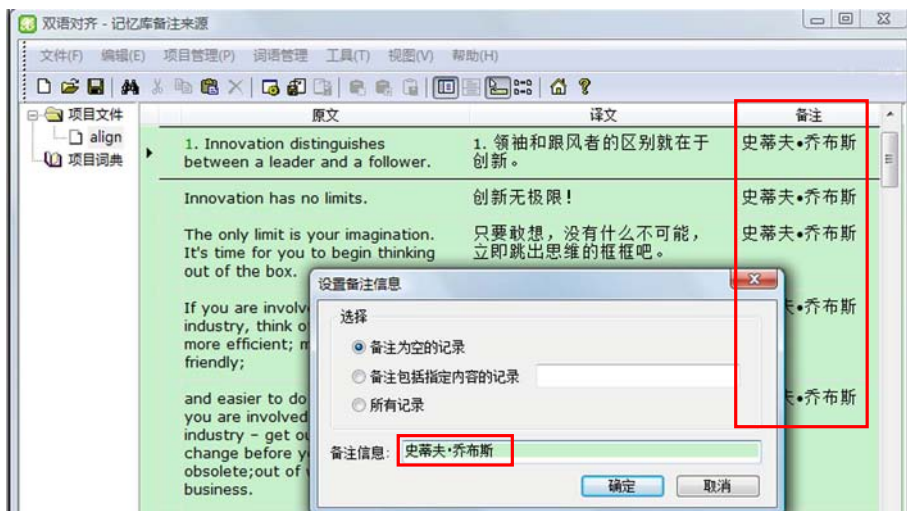
点击【工具菜单→设置备注信息】可以整批设置备注信息，该功能应用广泛，既可以 对翻译的句子按条件筛选设置备注信息，也可以对记忆库、词典等等整批设置备注信息， 注明它们的来源，为日后使用提供参考。以下将通过几个实例来说明它的应用。

例如：在预翻译时使用“在线自动”翻译的，在“备注”列中都会自动填其来源，如 果用 GOOGLE 翻译的，则对每一句会自动标注“GOOGLE”。此时若想删除这些备注信息 或者替换成其他的备注信息，就点击【工具菜单→设置备注信息】，如下图：在弹出的窗 口中选择筛选条件，选取需要替换的记录，并在“备注信息”中输入替换的内容。这里选 择了“所有记录”，“备注信息”中输入了空格，点击确定后就将所有记录的备注信息替换

为空格。其中“备注包括指定内容的记录”选项需要用户输入相应的筛选条件内容，例如输入“a”，则备注信息中包含“a”的记录全部会被选择上。



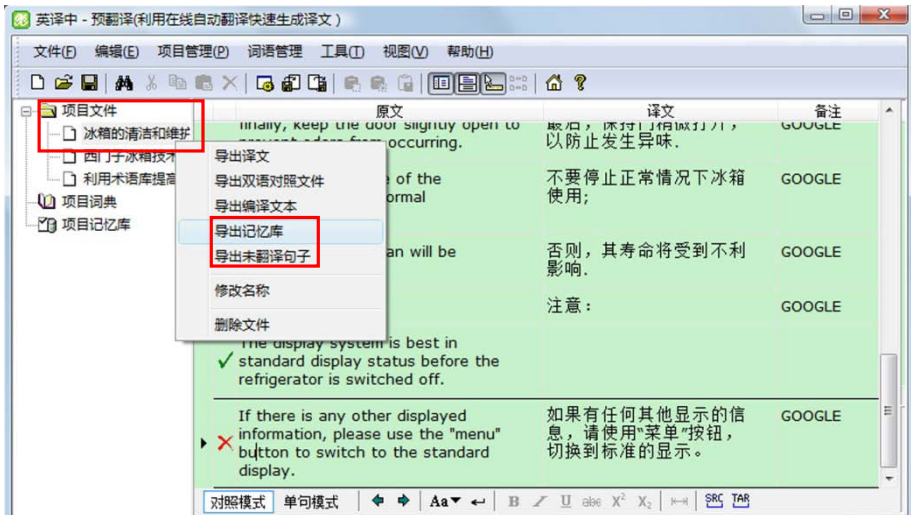
整批标注“备注信息”在创建记忆库时也是非常有用的，例如，在利用双语对齐功能建记忆库时，可以利用这个功能，对每句的出处进行备注，说明它的来源、分类等等，为日后的选用提供更多的参考信息。如下图中：



如果需要，也可以对“项目词典”中的词语进行整批标注备注信息，方法同上。

### 三、导出单个翻译文档的记忆库

雪人 CAT 的一个翻译“项目文件”下面支持导入多份需要翻译的文档，翻译后它们形成的记忆库全部汇集在“项目记忆库”中；另外，利用在线自动翻译的句子，如果没有经过任何的修改，是不会自动保存在“项目记忆库”中的。如果希望单独导出某份文档的记忆库（或将在在线自动翻译的句子直接导出记忆库），则需要在“项目文件”下面，点击某个具体的翻译文档，然后点击【右键菜单→导出记忆库】功能来实现。注意：如果利用【F4】标注为不正确的句子，导出记忆库时该句会被忽略。如下图：

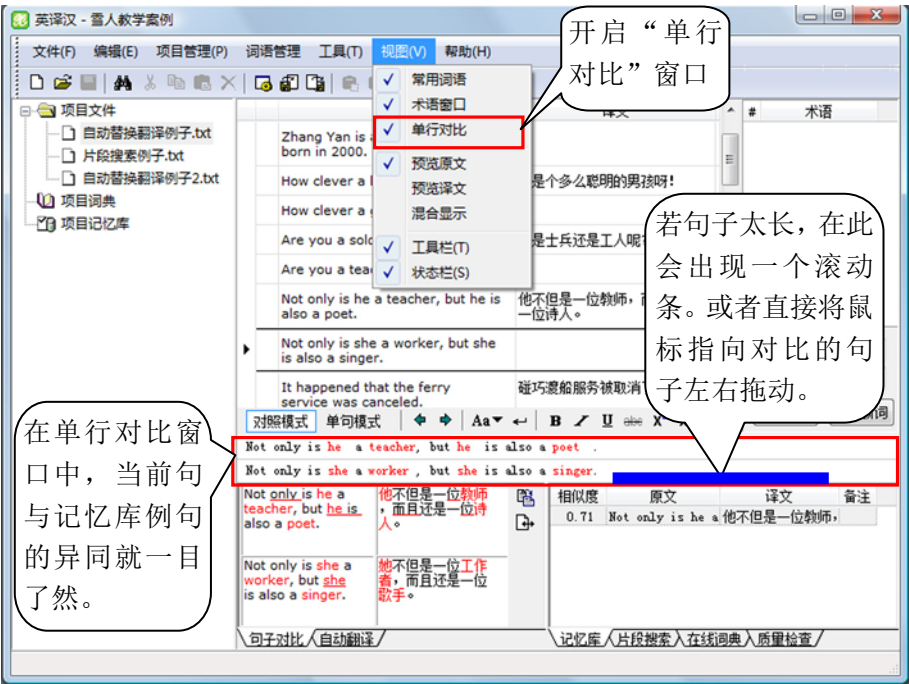




# 第六节 其他功能

## 一、单行对比清晰显示两句的异同

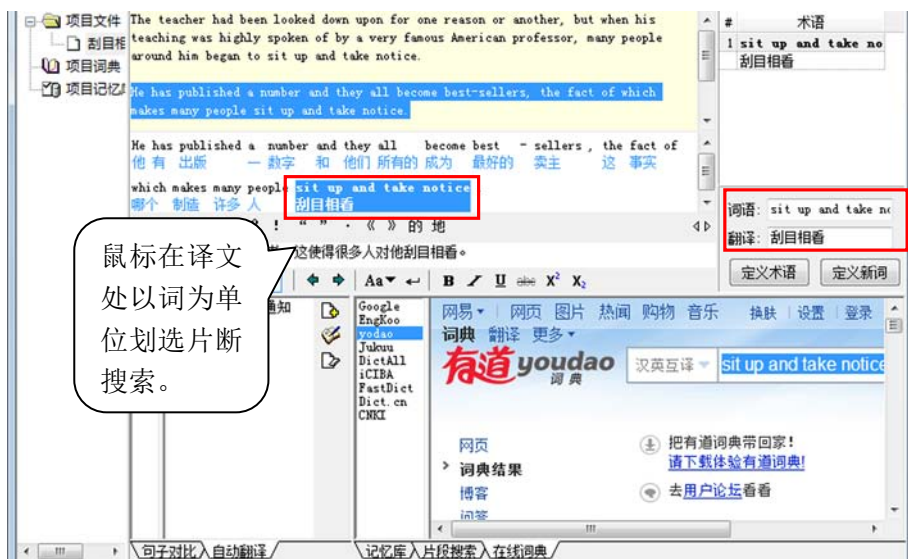
利用雪人的[单行对比]功能，可以非常方便地看到当前句与记忆库中的相似句子的异同，点击【视图→单行对比】菜单可以调出或关闭该功能，[单行对比]采用一一对应的显示方式，使得当前句与记忆库中的参考例句的异同一目了然。若句子太长，不能在一屏中完整显示，此时在对比窗口的底部，此会出现一个滚动条，如下图：为了节省屏幕空间，滚动条做得比较细小，若不注意，则不容易发现。另外，浏览长句时，也可以直接将鼠标指向对比的句子，按下鼠标左键，此时鼠标会变成一个“手掌”的形状，可以左右拖动句子来查看。



## 二、鼠标划选方式

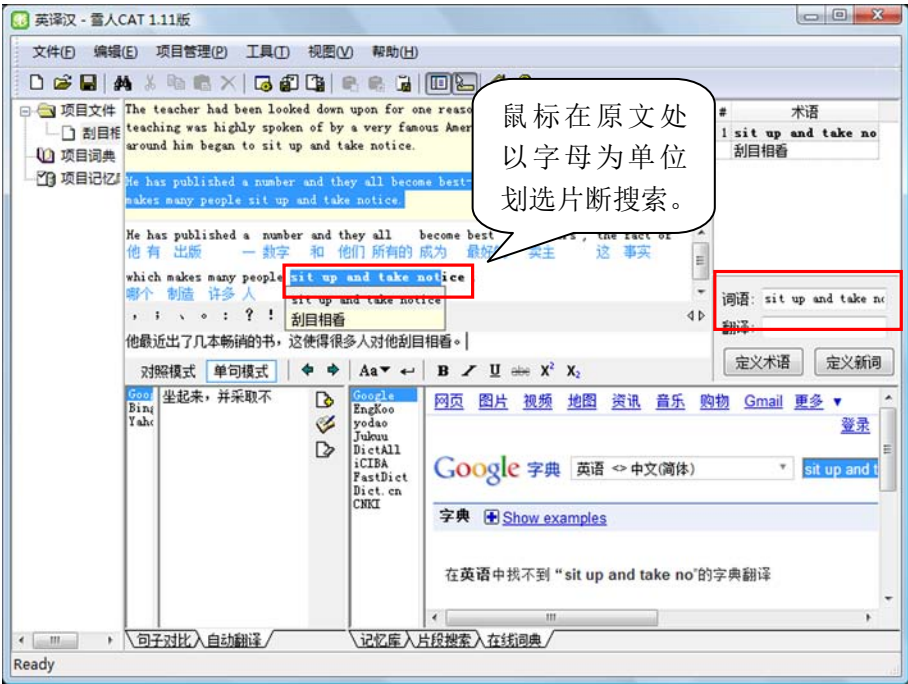
在单句输入模式中，当需要选择某个片段搜索时，软件有两种选词方式：一是以词为

单位的选择方式，即鼠标在**译文处划选**部分短语，这种方式选择快，它是以词语为单位选择移动的；如下图：鼠标在译文处划选了“sit up and take notice”并将它定义为“刮目相看”，则它的解释就会马上出现在屏幕取词窗口。



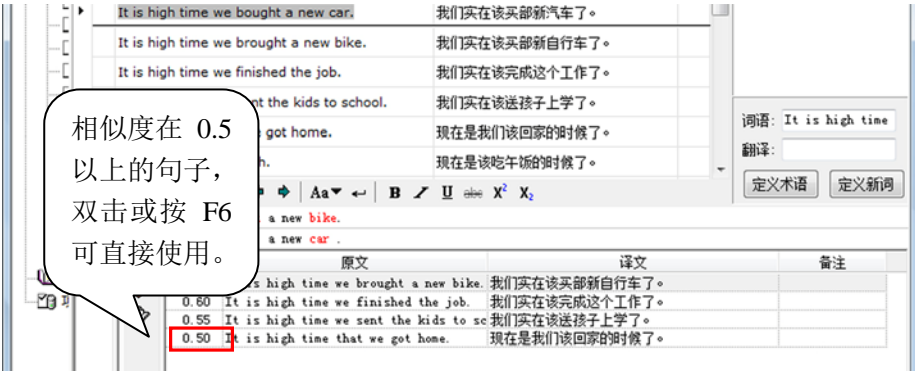
另一种选择方式是：鼠标**指向原文处划选**，这种方式以字母为单位移动，选择可以更加灵活，如下图：只划选了该短语的一部分。





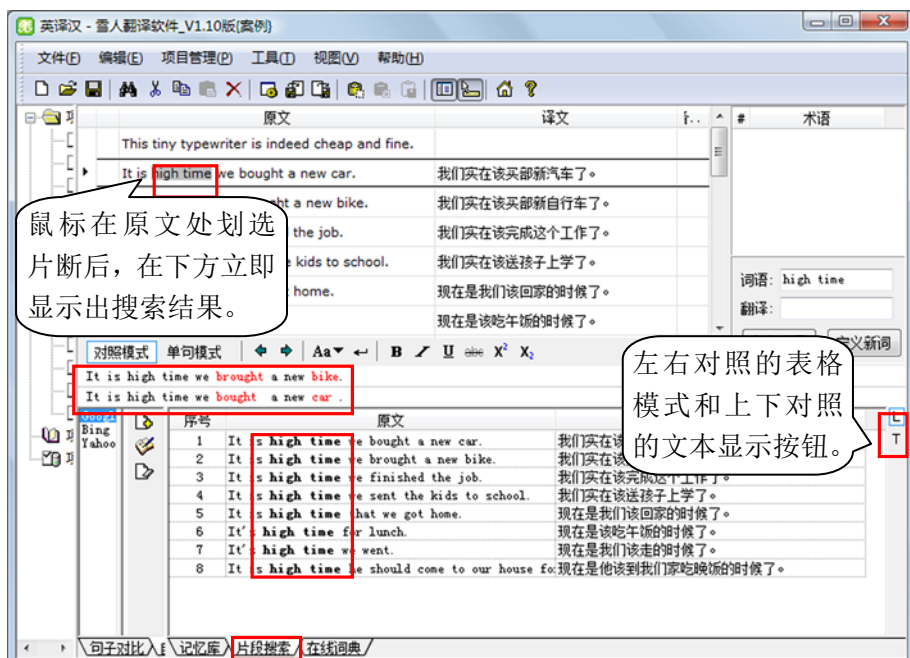
### 三、片段搜索的显示模式

对于相似度在 0.5 以上的句子，软件会自动显示在屏幕下方的记忆库页签中，双击或按【F6】键可直接在译文中输入，并对不同的部分自动做了相应的替换，如下图：

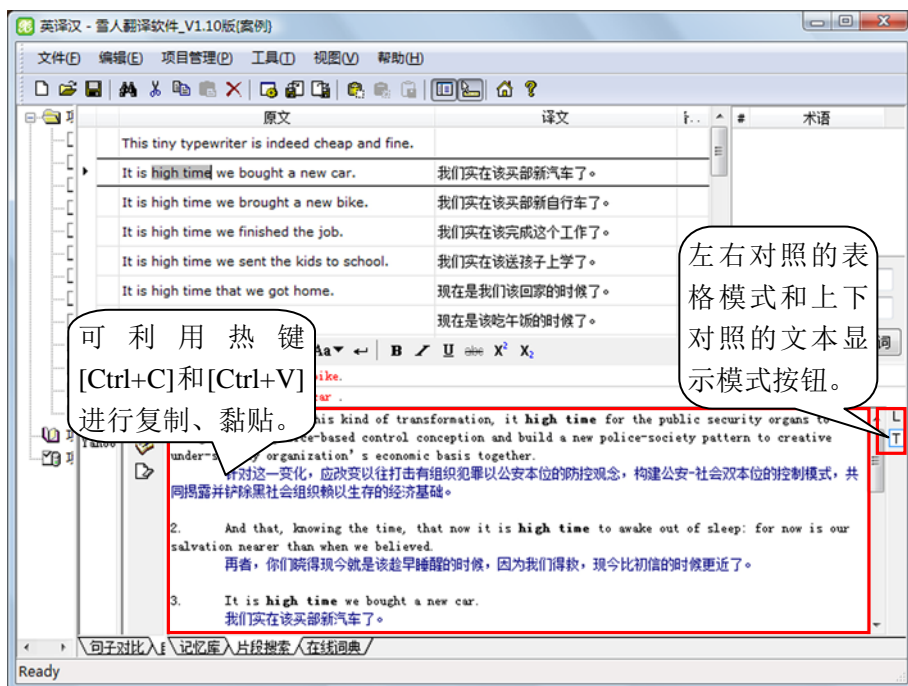


但是，由于能精确匹配的例句并不是很多，要能充分发挥记忆库的作用，建议使用软件强大的“片段搜索”功能。如下图：鼠标在原文处划选某片段后，在右下方的“片段搜索”窗口立即显示出搜索结果，软件默认以左右对照的表格模式来显示搜索结果，这种显示方式的优点是：在有限的窗口空间内显示更多的记录，缺点是：当遇到长句时不会自动

转行，查看不够方便，不能对查找的结果进行复制如下图：

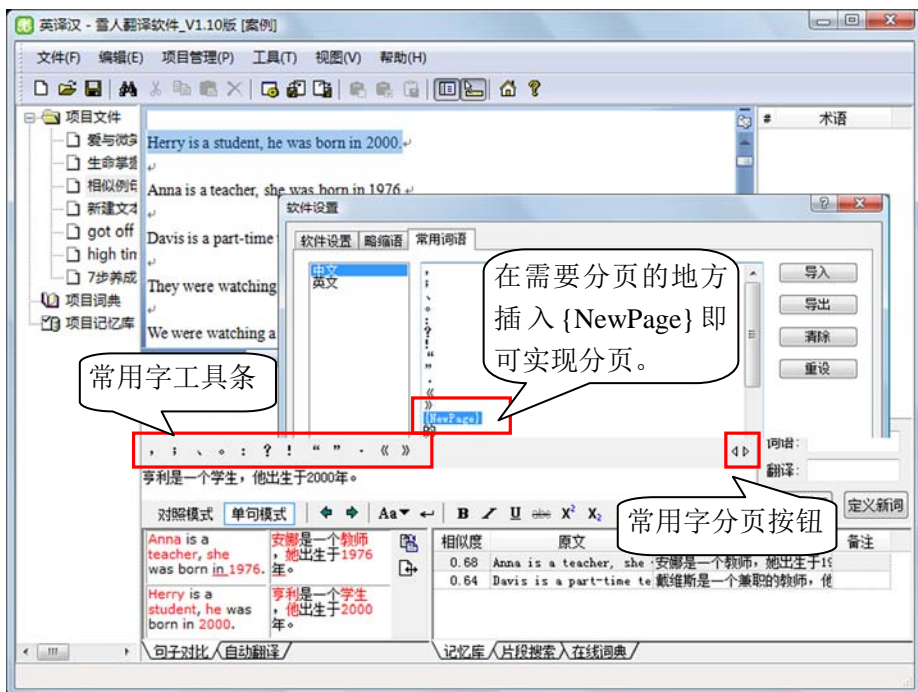


为此，软件增加了上下对照的文本显示模式，优点是可以自动转换，对长句查看很方便，而且可以将查找的结果复制到译文中或定义新词中（按热键【Ctrl+C】、【Ctrl+V】进行复制、粘贴）。两种显示方式可以通过点击该窗口右侧的“**L**”按钮（List：列表方式）和“**T**”（Text，文本方式）按钮随时切换，例如：当点击“T”按钮后，显示界面如下图：

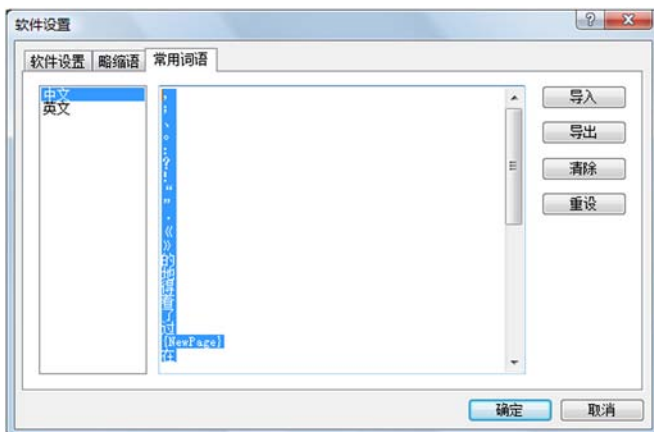


## 四、常用字自定义分页


为了加快译员的输入速度，雪人支持了常用字工具条，在【视图→常用词语】菜单可以调出或关闭，在【工具→软件设置】菜单的常用词语页签中可以定义常用字，如下图：如果定义的常用字过多，可以分类按页显示，这时在常用词语页签中需要分页的位置上插入 {NewPage} 即可实现分页显示，插入多个 {NewPage} 就可以划分为多页。



常用字工具条中显示的具体内容可以自己定义，定义的方法：点击【工具→软件设置】菜单，然后点击【常用词语】页签即可进行编辑、增减、导入等操作。如下图：



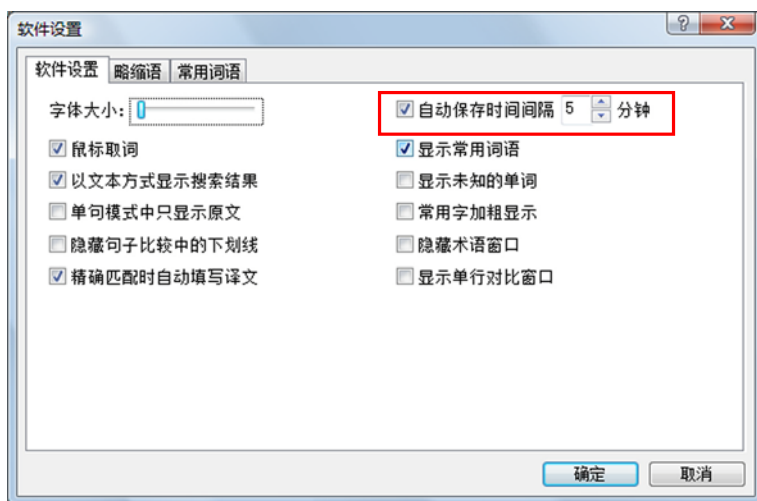
## 五、方便的查找功能

雪人翻译软件提供方便、快捷的查找功能，可以在原文查找某个词语或短语，也可在译文或备注中进行查找，单击工具栏的  查找按钮或按 **【Ctrl+F】**，可弹出查找窗口，如下图：按 **【F3】** 键可以查找下一个目标。



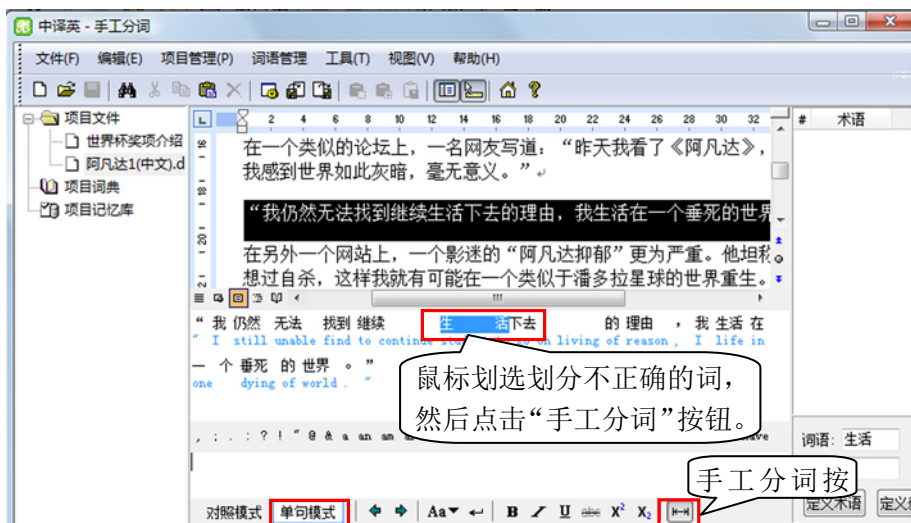
## 六、自动存盘

点击 **【工具→软件设置】** 菜单，可以看到这个功能，如下图：**该选项默认是不打开的，用户需要自行将该选项“打勾”，才能启动自动存盘功能。**自动存盘时间默认设置为每隔 5 分钟一次，注意这个时间间隔指的是从发生修改后开始计算的 5 分钟，若一直都没有更新，则不存盘。自动存盘不影响你当前的文件，而是自动另外生成一个相同文件名称，但后缀为 .SAV 的文件，若遇到断电、意外退出等情况的时候，将 .SAV 文件名的后缀删掉就可以直接打开，自动存盘是遇到意外情况的一种补救措施，所以用户还是应该向往常一样，该存盘时还是应该自行保存。**注意：新建的文件默认需要在用户进行第一次存盘后才能自动存盘。**

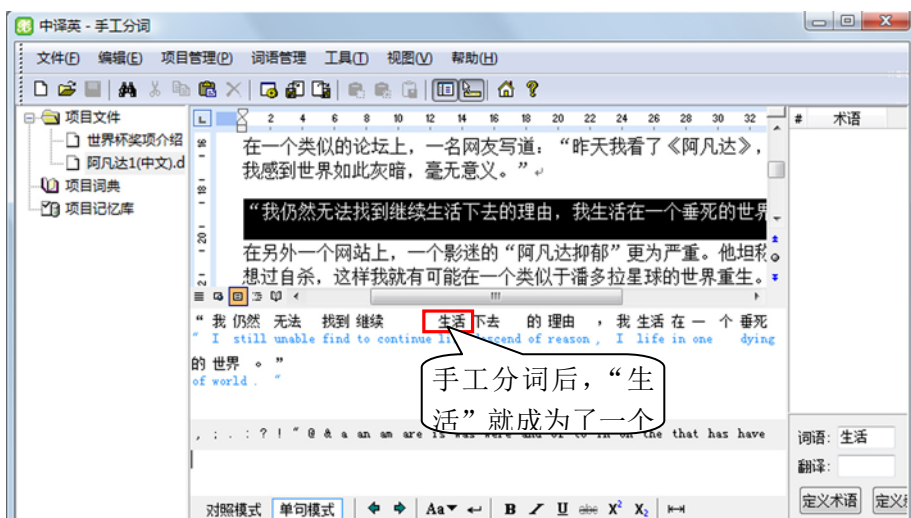


## 七、手工分词

有时软件自动的分词不一定正确，这时可以利用“手工分词”来进行修正。在“单句模式”下，若遇到自动分词不正确的情况，可以用鼠标划选划分不正确的词，然后点击“手工分词”按钮，就可以重新划分了。例如：下图中“生活下去”的划分不正确，



此时只要用鼠标划选“生活”二字，然后点击下方的“手工分词”按钮就将它断开成“生活 下去”了。如下图：

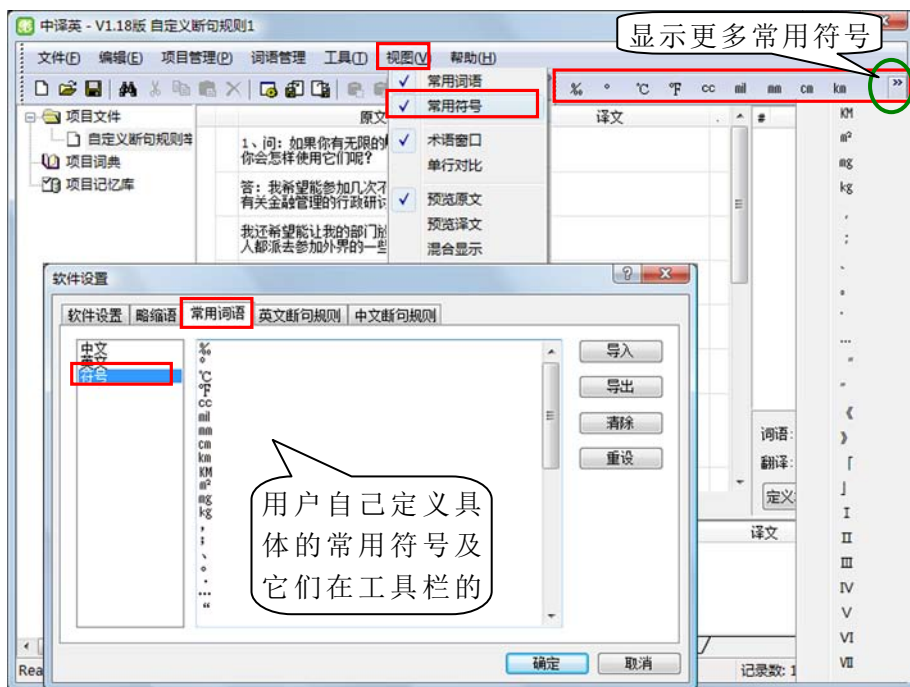


对于进行了手工分词操作的，目前版本不记录断词的设定，只用于当前句。



## 八、常用符号工具栏

翻译时，经常会遇到一些特殊符号的输入:例如%、℃、下等，为了提高输入速度，雪人增加了“常用符号”工具栏，用户点击工具栏中的符号即可输入，点击【视图→常用符号】菜单可以调出“常用符号”工具栏，如下图：当常用符号较多时，在右侧会出现一个“>>”按钮，点击它可以查看更多的常用符号。



工具栏中显示的“常用符号”的内容及显示顺序完全由用户决定，点击【工具→软件设置】菜单，在弹出的“软件设置”窗口中，点击常用词语页签，然后点击屏幕左侧的“符号”，就可以在此输入常用字或调整各个常用字在工具栏中的显示顺序。用户也可以利用 WORD 等工具将一些常用字编辑好，然后点击“导入”按钮整批导入，注意：“导入”新的常用符号会覆盖原有的常用符号，如果需要保存原有的常用符号，可以点击“导出”按钮，将它导出保存。

## 九、识别大写词语选项

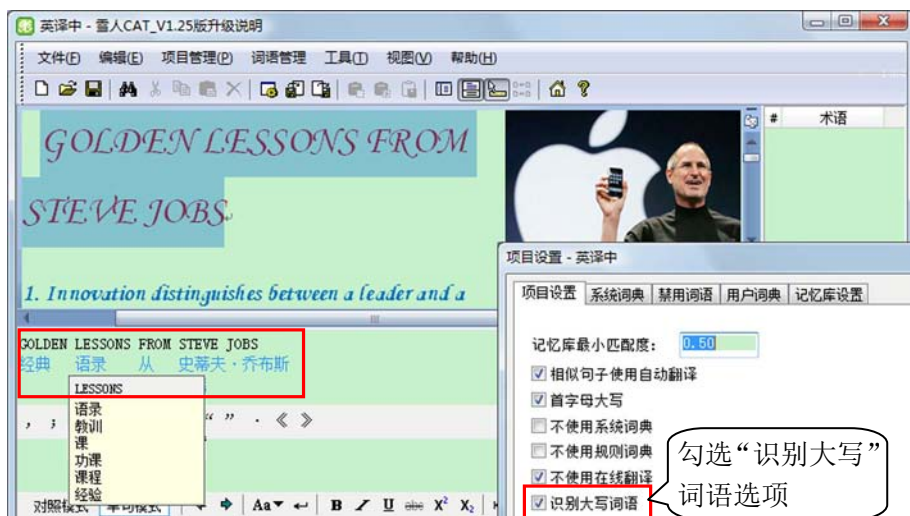
由于有些词语的大写和小写意思可能完全不同（如 it IT），所以雪人 CAT 的之前版本没有对大写的词语进行处理，遇到大写的词语会认为是生词，当鼠标指向该词时不会弹



出“屏幕取词”窗口，在词语自动翻译区亦不会自动翻译，如下图：

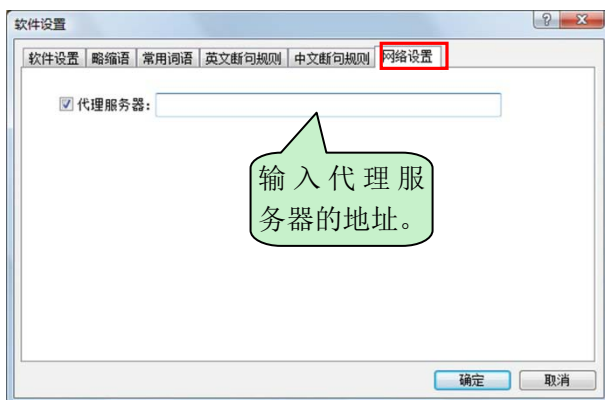


在 V1.25 版中，增加了“识别大写词语”选项，可以解决大写词语不能识别的问题。点击【项目管理菜单→项目设置】，勾选“识别大写”词语选项，此时可以看到在“词语自动翻译”区中出现了译文，当鼠标指向某个词语时，会弹出“屏幕取词”窗口，如下图：用户可以根据需要随时启用或关闭该选项。



## 十、支持使用代理服务器

由于有些用户的机器无法直接联网，而是通过代理服务器对外联网，在 V1.27 版中支持了这部分用户的需求，在【工具→软件设置→网络设置】中输入代理服务器的地址即可对外联网使用。



## 第四章 翻译各种文件格式

### 第一节 翻译 WORD 文档

雪人对 WORD 文档的支持与一般的 CAT 软件不同，雪人是在软件中直接嵌入 WORD，不用另外开启 WORD 的窗口，一个项目文件下面可以支持多份 WORD 文档。而一般的 CAT 软件是通过 WORD 的宏命令来实现的，它必须启动一个 WORD 窗口，而且同时只能有一份 WORD 文档，若打开多份 WORD 文档时，会出现混乱。翻译软件对文档中的字体样式、文本绕图排版、上下标等都有很好的支持，而且还支持以字符形式插入的图片（如“国庆60周年”中的 60）；雪人对表格支持也很好，翻译时只需专注表格的内容，表格样式交由软件自动处理。

雪人的样式控制码很少，相当于同类软件样式码的十分之一，这极大地减轻译员的工作量。在雪人中，很多样式由软件自动处理了，当一句中只有一种样式时，根本不出现样式码，只要当一句中有两种以上的字体样式时，才会出现样式码。如果不想在翻译中处理样式，则可以不理会那些样式码，即使翻译时忽略所有的样式代码，雪人也可以导出一个样式基本正确的译文，导出后再进行样式排版处理。雪人可以随时导出译文，而有些 CAT 软件则必须要求用户一个不漏地完全正确处理完所有的样式码后，才可以导出译文。

下面我们将通过几个事例来讲解雪人对 WORD 的出色表现：

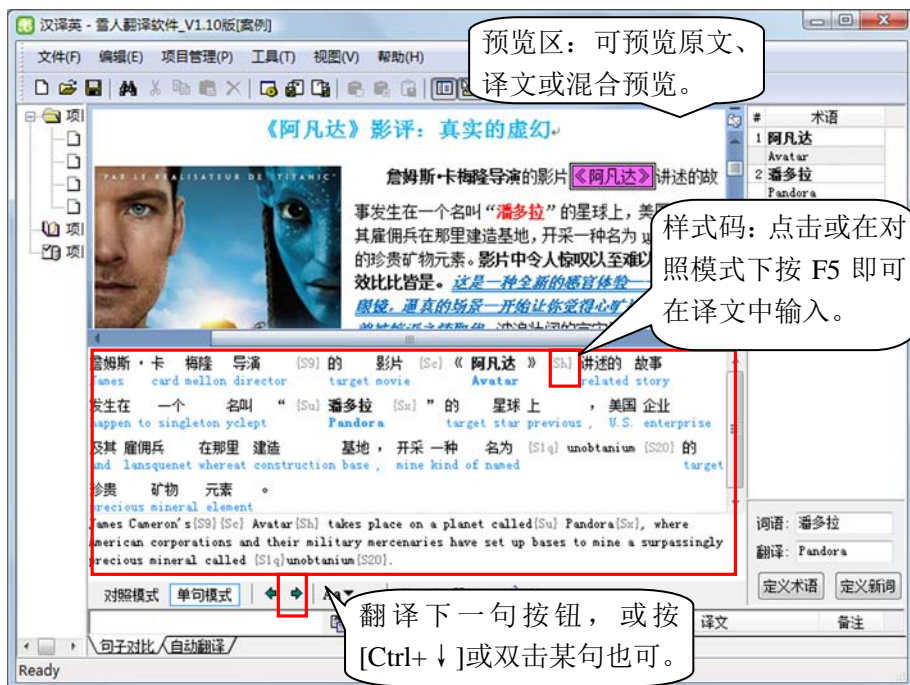
#### 一、对图文并茂 WORD 文档的支持

我们在安装有 WORD2007 版本的机器上导入一个关于“《阿凡达》影评”的 WORD 文档，导入后如下图：我们看到该文档有插入绕图，字体样式也比较多，具体有：

1. 标题：“《阿凡达》影评：真实的虚幻”浅蓝色、加粗、四号字体
2. 第一句：“詹姆斯·卡梅隆导演的影片《阿凡达》讲述的故事发生在一个名叫“潘多拉”的星球上，美国企业及其雇佣兵在那里建造基地，开采一种名为unobtainium的珍贵矿物元素。“该句有 5 种样式：（1）“詹姆斯·卡梅隆导演”加粗；（2）《阿凡达》加了底纹及边框；（3）“潘多拉”红色加粗；（4）unobtainium加了下划线；（5）其他的为普通五号字体。
3. 第二句：“影片中令人惊叹以至难以置信的特效比比皆是。”该句全句样式都相同，即全部加粗。
4. 第三句：“这是一种全新的感官体验——戴上 3D 眼镜，逼真的场景一开始让你觉得心旷神怡，但很快就被惊讶之情取代。”该句全句样式都相同，即全部加粗、斜体、

蓝色、加双线下划线。

进行了上述字体样式分析后，我们需要提取这些字体的样式。雪人提取字体样式的方式有两种：一种是边翻译边自动提取，翻译到哪句时马上自动提取该句的样式。由于提取字体的样式需要进行大量的计算及字体变化的比对，若样式复杂，对配置不高的机器当第一次双击某句进行翻译时，会感觉到有 1~2 秒的等待，样式一经提取后会自动保存，所以光标第二次回到该句时反映就很迅速了。另一种是首先一次性地提取所有字体样式，这样当翻译时就不存在任何的等待，点击【工具菜单→预提取字体样式】功能即可，不过这有时需要花费一定的时间来计算字体的样式，快则几秒钟，慢则几分钟，视文章的大小、样式变化的复杂度而定，预提取字体样式的进程可以随时中断，中断后没有提取的那部分会在翻译时逐句自动提取。预提取字体样式功能对 WORD2007 版本才有效，对 WORD2003 版本无效；该功能只用于提取字体的样式，对于图片等其他样式软件在导入时就自动处理。



提取完字体样式后，我们看到雪人的样式控制码很少，如果整句的样式相同，由软件自动处理，根本不会出现样式码；当在一句中有多种不同的字体样式时才会出现样式码。如上图：标题“《阿凡达》影评：真实的虚幻”这几个字的样式为：浅蓝色、加粗、四号字体。但是由于整句都相同，所以根本不出现样式码，译员只需关注该如何翻译就可以了。第一句有 5 种样式，我们从上图中看到软件出现了多个“{}”灰色的样式码。所以，

对于这一句我们翻译完成后，还需要在译文的相应位置插入回样式码，点击上面的样式码或在对照模式下按【F5】键即可输入。译好该句后，按【Ctrl+↓】（或双击下一句、或按下方下一句按钮）进行下一句的翻译。上图中我们选择的是预览原文的方式，点击【视图菜单→预览译文】，我们看到刚才已经翻译好的标题及第一句的效果，如下图。

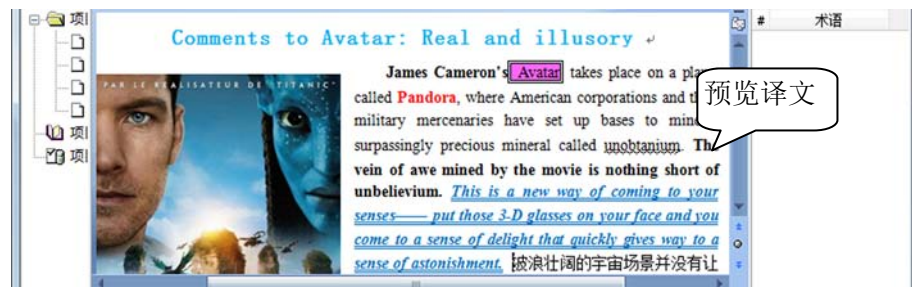
接下来我们翻译第二句：“影片中令人惊叹以至难以置信的特效比比皆是。”由于该句全句样式都相同，即全部加粗，所以软件中没有出现样式码，样式由软件自动处理。



跟着翻译第三句：“这是一种全新的感官体验——戴上 3D 眼镜，逼真的场景一开始让你觉得心旷神怡，但很快就被惊讶之情取代。”该句全句样式也都相同，即全部加粗、蓝色、加双线下划线。同样软件自动处理了它的样式，不出现样式码，译员只需关注译文本身，样式码少可以极大地减少译员的工作量。

另外，如果在单句模式下，全选后按【Ctrl+T】可以将“{}”里面的内容变成样式代码，否则“{}”里面的内容不认为是样式代码。

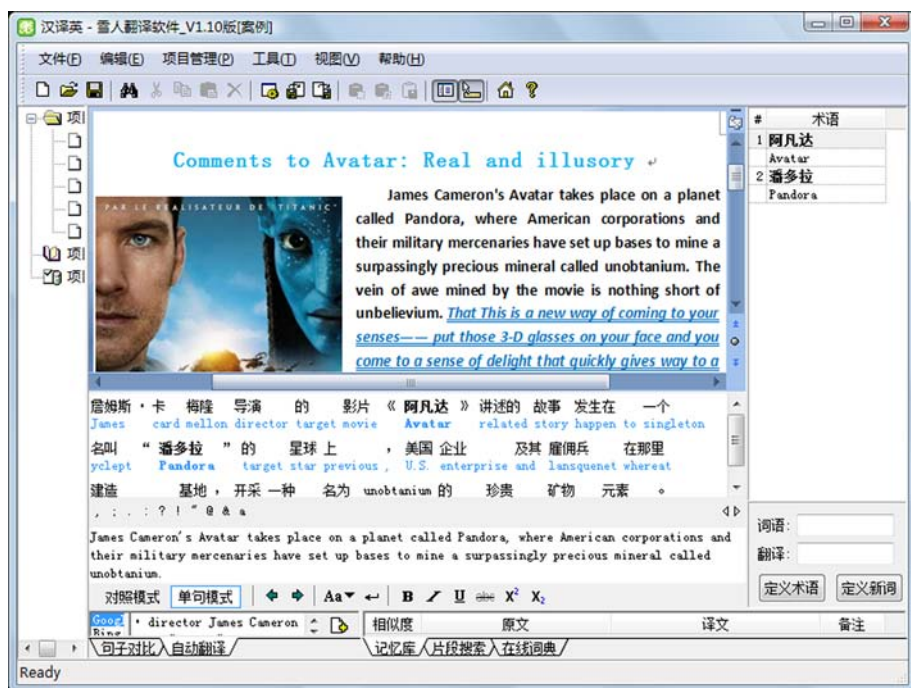
雪人的译文实时预览技术，在翻译时会给译员带来更多的方便，译好一句后立刻看到译文的效果，译文随时自动更新，随翻译随预览，这点非常方便实用，译文预览效果如下图；不像有些翻译软件，必须点击刷新按钮后，才能预览。





不过,上述雪人软件的近乎完美的表现需要用户安装有 WORD2007 版本,对于 WORD2003 版本,由于微软有部分样式接口不开放,所以效果会有一定的影响。【工具菜单→预提取字体样式】功能只用于提取字体的样式,不涉及表格、绕图等方面,该功能只对 WORD2007 版本才有效,对 WORD2003 版本无效。对于 WORD2003 版本,可以处理一些简单的字体样式,即整句中只有一种样式的情况,但对于“单句”中有两种以上的样式会统一按句首的样式处理,这需要导出译文后进行修正。其他如表格、绕图等等的排版格式在 WORD2003 版本也可以很好地支持。

下图中是在 WORD2003 版本中进行导入后翻译的效果:针对上例,第一句的字体则全部按句首加粗处理了,其他句的字体样式由于单句中只有一种,软件自动处理还是正确的。

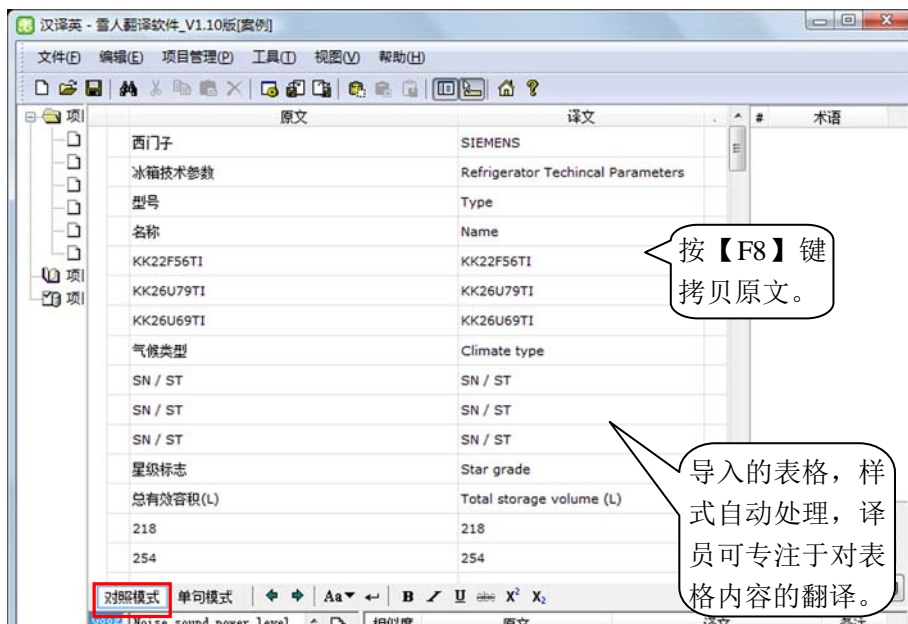


如果原文的样式不是太复杂,字体不会出现“一句”中变化多种样式的话,WORD2003 的输出效果也还可以。如果字体样式复杂或对字体样式有精确的要求,建议还是安装一个 WORD2007 版本。WORD2007 版本可以直接导入 DOCX 格式的文件,WORD2003 需要从光盘的“TOOL”文件夹中安装一个微软的 DOCX 格式的支持文件后,才能将 DOCX 格式的文件导入到雪人中。

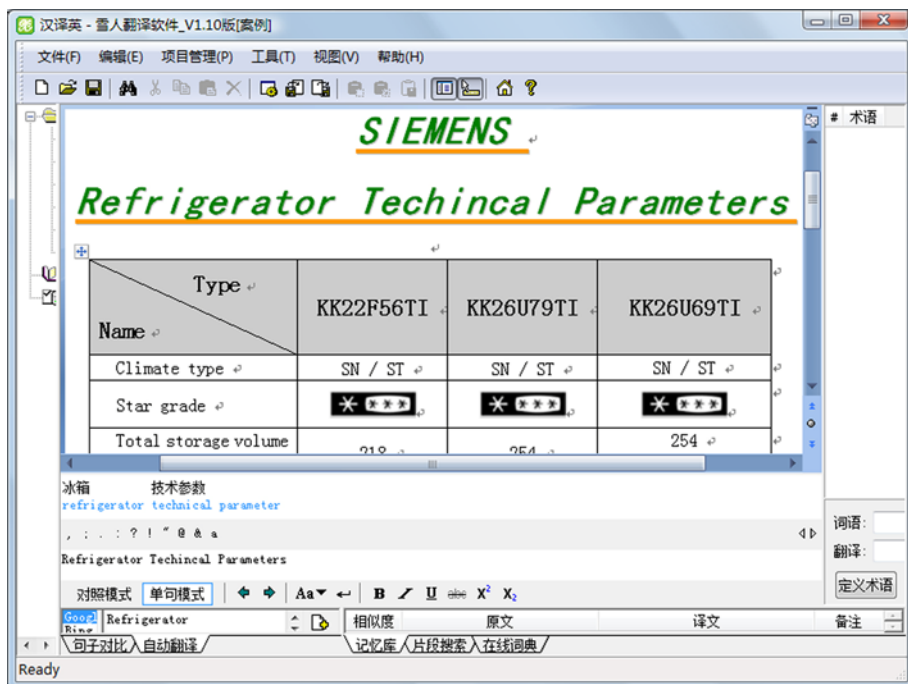





以将原文快速拷贝过来。



当然，也可随时切换到单句模式中，因为单句模式可以同时预览原文或译文，如下图：



待将表格所有内容翻译好后，点击工具栏的导出按钮或右键菜单的导出功能，将译好的表格导出。打开刚才导出的表格，我们发现表格所有的样式与原文保持高度一致。如下图：

12122123124125126127128129130131132

SIEMENS

Refrigerator Techincal Parameters

Type Name	KK22F56TI	KK26U79TI	KK26U69TI
Climate type	SN / ST	SN / ST	SN / ST
Star grade	***	***	***
Total storage volume (L)	218	254	254
Freezer compartment storage volume (L)	51	64	64
Freezer compartment temperature (℃)	-18	-18	-18
Power supply	220V 50Hz	220V 50Hz	220V 50Hz
Freezing capacity (kg/24h)	17	10	10
Power consumption (kW·h/24h)	0.66	0.78	0.76
China energy level	1	1	1
Refrigerant	R600a ( 48g )	R600a ( 41g )	R600a ( 41g )
External dimensions (mm)	590x649x1750	640x633x1850	640x641x1900
Rated current (A)	0.6	0.8	0.8
		38.9 for normal operation (42 for	38.9 for normal operation (42 for

III

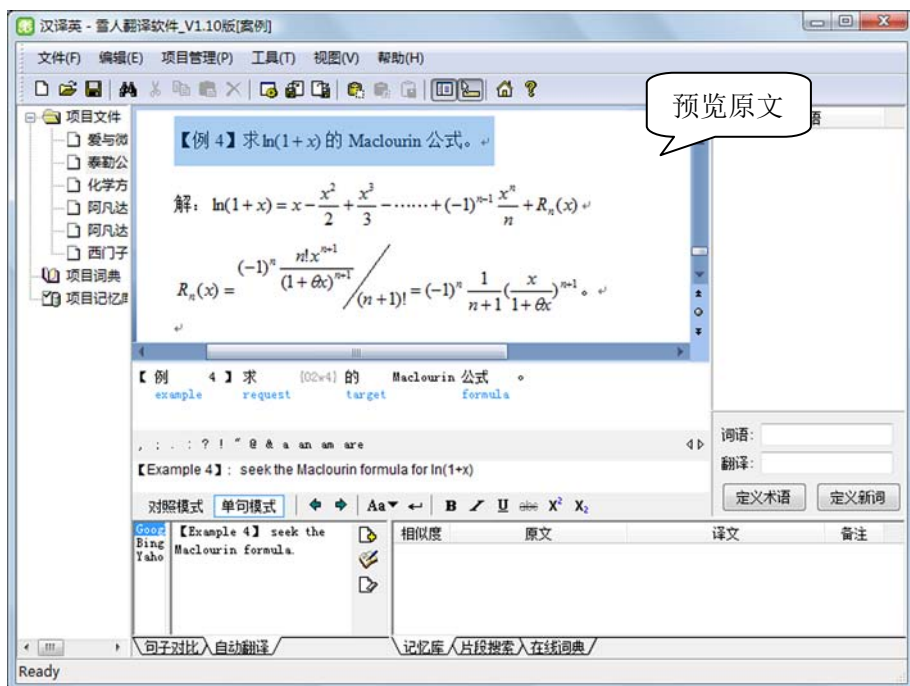
使用雪人 CAT 软件进行高效翻译的其中一个原因是：**雪人断句准确和样式码少，绝大多数的样式码雪人可以自动处理，根本不出现在屏幕中。**同样一篇文章它的样式码在其他翻译软件中有近百个，而在雪人中可能只有十来个，要将这近百个的样式码在译文中准确插入得花费不少功夫。如上述的表格案例，原文虽然有不少样式，但导入到雪人软件中没有出现一个样式码，所有的样式软件自动处理了，译员只要集中精力翻译表格内容就行了；而同样的表格导入到其他同类软件中，样式码就有 35 个。

有时，译员可能会在翻译时不处理样式，待导出译文后再进行样式处理。雪人可以随时导出，即使没有译完或在译文中忽略部分（或全部）的样式，一样可以导出。而有些翻译软件则要求用户必须一个不漏地插入所有的样式码后，才可以导出译文，否则无法导出！

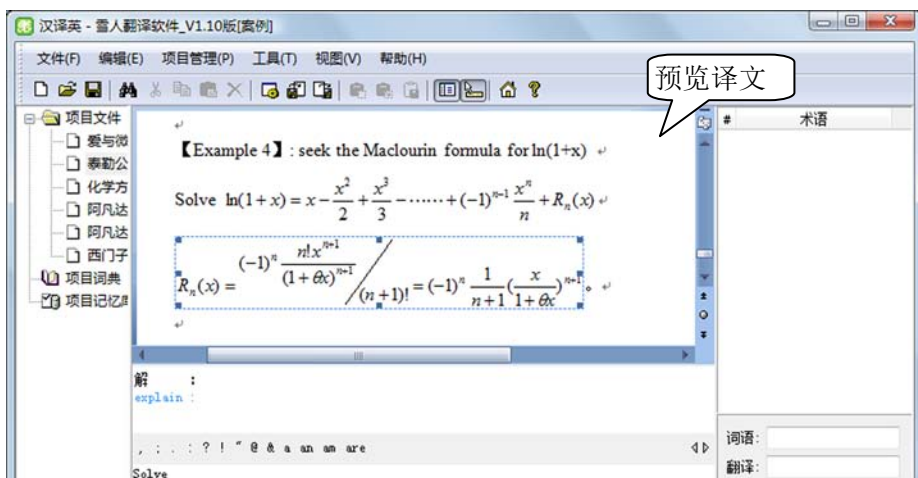
三、对复杂数学公式的支持

对于遇到样式复杂的公式，若使用**对照模式**翻译，可能会只看一些样式码，往往不

知所云；但此时使用单句模式翻译就显得非常直观方便，这就是雪人两种翻译模式并行的优势所在了！如下图：导入了一个“泰勒公式”文档，翻译“求  $\ln(1+x)$  的 Maclourin 公式”，我们看到公式是比较复杂的，有指数形式、有多种数学符号，分子式等等。



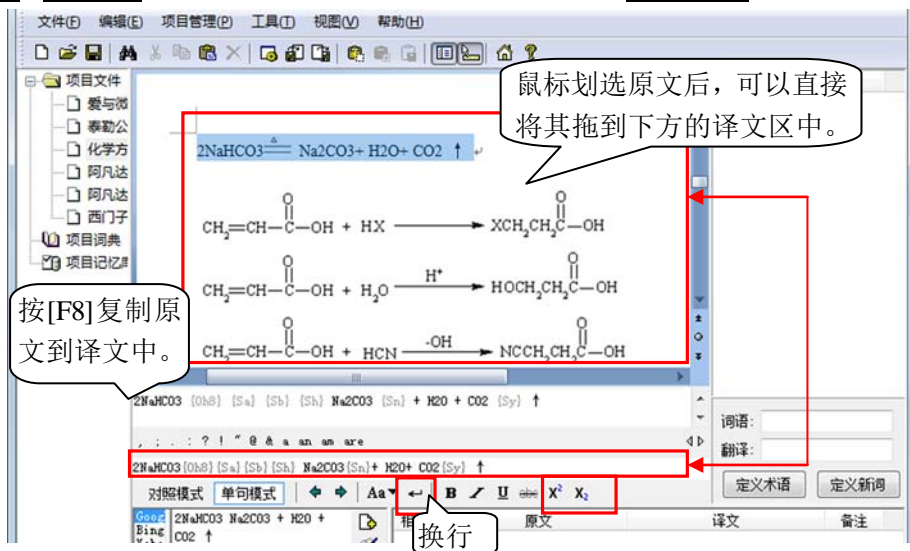
但使用雪人翻译却相当简单，只要翻译“求  $\ln(1+x)$  的 Maclourin 公式”这几个字即可，其他的根本不用处理，软件自动转换了，甚至连样式码都没有出现。如下译文预览图：



## 四、对复杂化学式的支持

化学式会经常涉及到上下标，雪人可以很好地处理上下标的问题，如果是 WORD2007 版本，上下标的样式会自动处理，在“单句模式”下不出现其样式码。如果是 WORD2003 版本，会忽略了上下标的样式，此时可以利用屏幕下方的上下标工具按钮来标注。另外，在[单句模式]中，雪人还有一个非常方便之处就是：在预览区中，用鼠标划选某些内容后，直接将其拖到译文输入区中，其样式也会一同被复制下来，这个对局部复制非常实用。



如果译员需要对公式进行换行，可以点击[换行]按钮，用于在译文中加入换行的支持，译员可根据自己的喜好对译文进行换行。如果译员需要对译文进行转换大小写，可以按[大小写转换]按钮，选中需要转换的译文，单击[大小写转换]按钮，便会有[首字母大写]、[全部大写]和[全部小写]的选择，在【项目管理→项目设置】中，也有[首字母大写]的默认设置。



### 注意事项：




WORD 文档的导入及预提取字体样式需要一定的时间，建议单份 WORD 不要太大，以 30～50 页为宜，对于大型的待翻译 WORD 文档，建议拆分成多个小的文档导入到雪人。为了保证译文样式导出的正确性，请保持导入、导出时 WORD 版本的前后一致性，即：导入时若是在 WORD2003 版本的机器上进行了，那么导出译文时也应该在装有 WORD2003 版本的机器上进行导出；若导入时是在 WORD2007 版本的机器上进行了，那么导出译文时也应该用 WORD2007 版本。





## 五、隐藏预览区

“单句模式”默认状态下，在屏幕的中上部是一个比较大的预览窗口，由于屏幕空间有限，如果遇到特别长的句子，下面的翻译区就显得空间不够了，此时可以点击工具栏的隐藏预览模式按钮，将预览区隐藏起来，使翻译区拥有更大的编辑区域。如下图：再次点击按钮，可以切换回预览窗口。



## 六、复制原文及组合译文

在屏幕中部的工具栏中，增加  和  两个按钮，其中  按钮用于复制所选择部分的原文，或者用鼠标选择需要的原文后按住左键不松手拖动到目标位置上也可以实现复制原文的目的。热键【F8】是复制整句的原文。

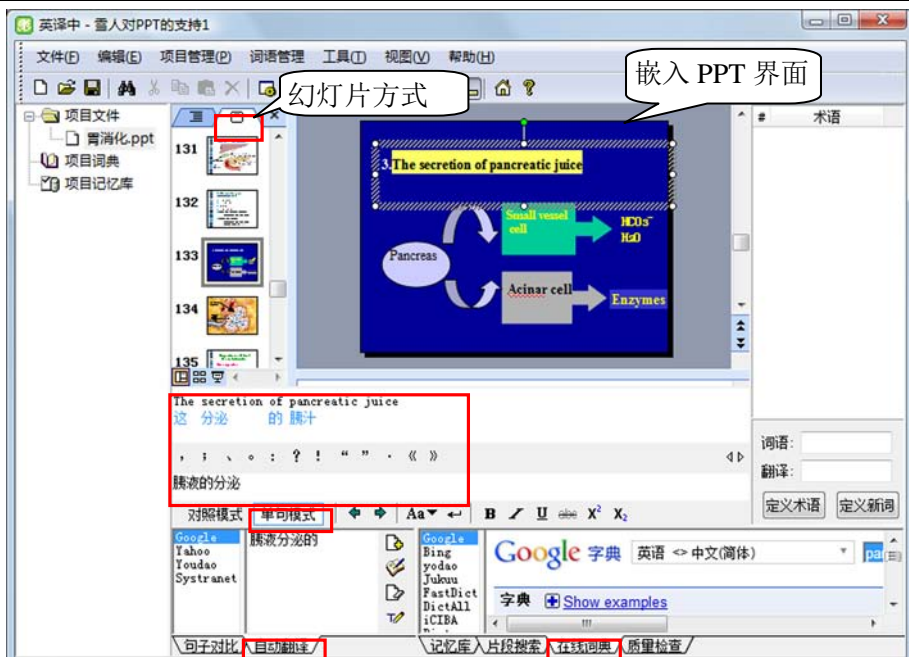
 按钮用于将自动提示的词语翻译组合成译文填写在译文区中，当译文为空时，点击  按钮，即可将词语翻译组合成句子填写在译文中，这在“对照模式”和“单句模式”下都支持，当译文中已经有内容时，点击  按钮则不起作用，如下图：另外，在“单句模式”下，还可以在“词语提示区”中用鼠标划选部分的译文后，点击  按钮来填写所划选部分的译文，不过，部分填写译文功能在“对照模式”下则不支持。



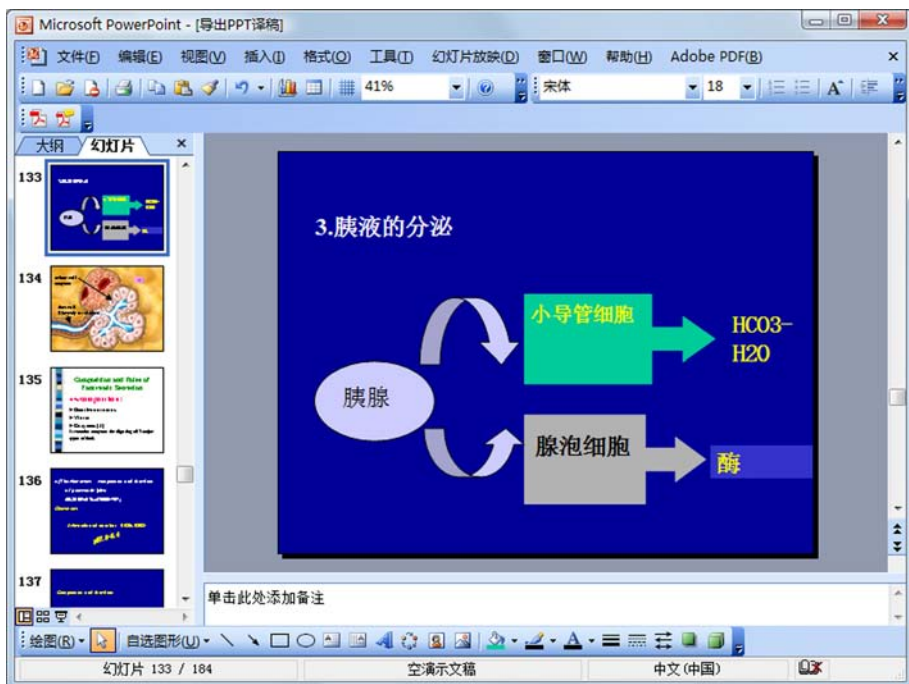


## 第二节 翻译 PPT 文档

将 PPT 格式的文件导入到雪人 CAT 的翻译项目中，对翻译 PPT 文档，“单句模式”可以直接预览 PPT 的原文，会比“对照模式”直观、方便。如下图：在“幻灯片”浏览方式下（不要在大纲模式下），单击需要翻译的句子，即可开始翻译。译者只需要关注翻译本身，而不必理会 PPT 的样式、排版等问题。



翻译好后，将译文导出，即可生产与原稿一样的的 PPT 译文。导出的译稿如下图：



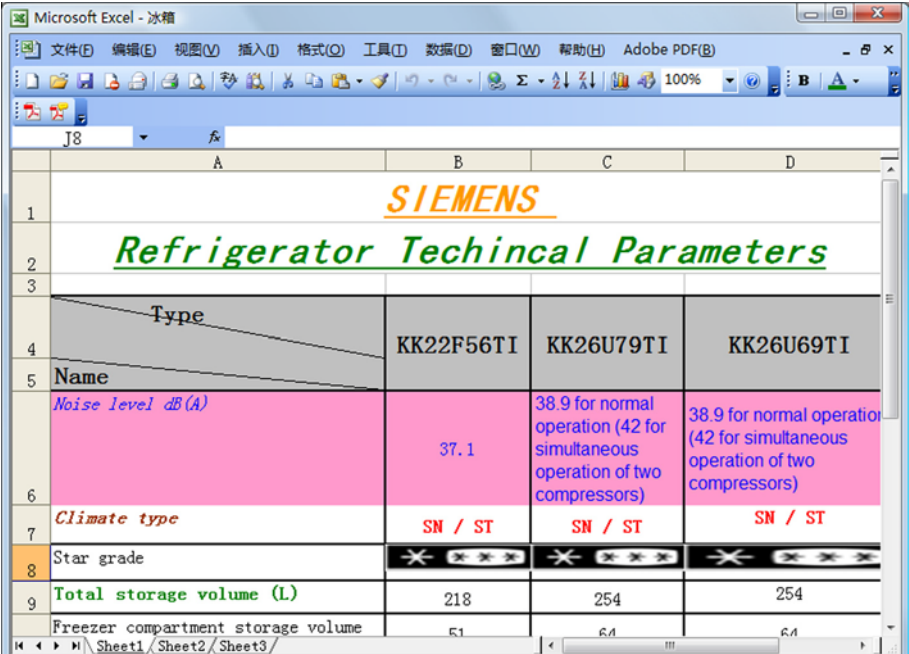


在翻译 PPT 时几点需要注意的地方：

- 1. 避免修改原文，修改后的内容不会存盘，同时会造成定位不准确，可以重新打开当前文件或切换项目中的文件进行恢复原文。
- 2. 导出译文后，部分没有生成的缩略图可能没有正确显示，可以切换到大纲模式再选择缩略图进行刷新。
- 3. 在 Power Point 2003 中，直接选择表格中的文字时不能自动定位到相应的句子。
- 4. 组合后的文字不能直接翻译，需要预先进行拆分。
- 5. PPT 预览无法在大纲模式下定位句子。

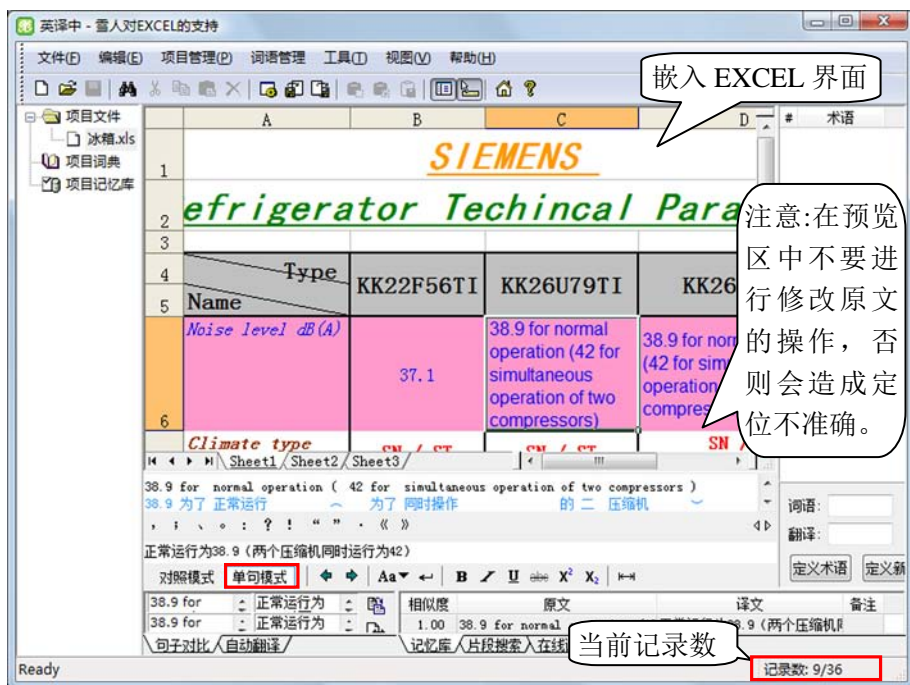
第三节 翻译 EXCEL 文档

支持将 EXCEL 嵌入雪人翻译软件中，例如我们欲翻译的 EXCEL 文档如下图：格式比较复杂，有图片，字体样式及颜色等的变化较多。



SIEMENS			
Refrigerator Technical Parameters			
Type	KK22F56TI	KK26U79TI	KK26U69TI
Name			
Noise level dB(A)	37.1	38.9 for normal operation (42 for simultaneous operation of two compressors)	38.9 for normal operation (42 for simultaneous operation of two compressors)
Climate type	SN / ST	SN / ST	SN / ST
Star grade	✖☆☆☆	✖☆☆☆	✖☆☆☆
Total storage volume (L)	218	254	254
Freezer compartment storage volume	51	64	64

我们将上面的 EXCEL 文件导入雪人，导入后在“单句模式”下可以预览原文，如下图：



译员翻译时，只需关注译文本身，不必理会字体的样式、颜色、图片等等。翻译完成后，在屏幕左侧点击【右键菜单→导出译文】功能，可以看到雪人生成的译稿，可以严格保持与原文的一致性。如下图：

Microsoft Excel - 导出译文——冰箱

文件(F) 编辑(E) 视图(V) 插入(I) 格式(O) 工具(T) 数据(D) 窗口(W) 帮助(H) Adobe PDF(B)

100%

西门子  
冰箱技术参数

	A	B	C	D
1				
2				
3				
4	型号	KK22F56TI	KK26U79TI	KK26U69TI
5	名称			
6	噪声功率级dB(A)	37.1	正常运行为38.9 (两个压缩机同时 运行为42)	正常运行为38.9(两个压 缩机同时运行为42)正 常运行为38.9(两个压 缩机同时运行为42)正 常运行为38.9(两个压 缩机同时运行为42)
7	气候类型	SN / ST	SN / ST	SN / ST
8	星级标志	★ ★ ★ ★	★ ★ ★ ★	★ ★ ★ ★
9	总有效容积(L)	218	254	254
10	冷冻室有效容积(L)	51	64	64
	冷冻室温度(℃)			

Sheet1 / Sheet2 / Sheet3 /

就绪 数字

注意：在雪人界面的预览区中不要进行修改原文的操作，即使修改了原文，亦不会保存修改的结果，而且还会有可能造成定位不准确的问题。若出现定位不准确的问题，通过重新打开文件或者切换到其他预览的项目再切换回来可以恢复。

## 第四节 利用剪切板翻译各种格式的文档

雪人有非常强大的剪切板功能，通过它，可以翻译网页、PDF、PPT 等等一切可以复制粘贴的文档。下面我们将举例介绍剪切板的具体应用：

### 一、翻译网页

通过剪切板我们可以很方便地对网页文件进行翻译，首先将要翻译的内容拷贝到剪切板中，然后将其粘贴到雪人软件，翻译好后，利用雪人的“**导出剪切板**”按钮，将译文导出，将其粘贴到网页的原来位置上就可以了，译员只需关注翻译本身，而无需太多地去关注网站的排版、样式等等。例如我们要翻译的网页如下图：




我们首先在网页编辑软件“Dreamweaver”中复制“雪人计算机辅助翻译”这个标题到剪切板中，然后切换到雪人翻译软件，在雪人软件中既可以以剪切板的内容自动新建一个文件，亦可以将剪切板的内容粘贴在当前文件的后面，这里选择工具栏的“**粘贴为新文件**”按钮，以剪切板的内容自动新建一个文件，如下图：则刚才复制的“佛山市雪人计算机公司”这几个字已经粘贴到雪人软件中，并且在左侧的“项目文件”夹中自动增加了一份名称为“clipborad”的文件，将此句翻译好后，点击工具栏的“**导出到剪切板**”

按钮将译文导出。



然后切换回网页编辑软件“Dreamweaver”中 按“粘贴”键即可将刚才的译文粘贴到原处即可。

工具栏的为“粘贴内容”按钮，它是将剪贴板的内容粘贴到当前文件的后面，即在左侧不增加一个“clipborad”的文件，而是粘贴在当前文件后面。

## 二、翻译 PDF 文件

以上是讲述如何利用剪切板来对网站进行翻译，其实剪贴板功能是相当灵活的，通过剪贴板的功能，可以实现对各种支持编辑的文件格式的翻译，例如对 PDF 格式文件的翻译，在 PDF 编辑器中将要翻译的内容复制到剪切板中，然后将它粘贴到雪人翻译软件中进行翻译，翻译好后将译文导出到剪切板后，按“粘贴”按钮就可将译文粘贴回原处了。如图，在 PDF 编辑器中复制一段内容到剪切板。

然后将所复制的内容粘贴到雪人翻译软件中进行翻译，译好后导出到剪切板，然后在 PDF 编辑器中按“粘贴”按钮就可以将译文粘贴回原处了，非常的方便！

### Operation



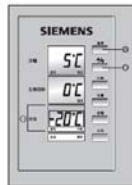
#### Adjusting the temperature in refrigerator compartment

- Press button 1 to activate Temperature Display (1).
- Press button 2 repeatedly until the desired temperature is displayed. It will decrease 1°C by each press and vary from 2°C to 10°C.
- Stop pressing it, the last selected value will be confirmed and memorized. Display (1) will show the selected temperature.



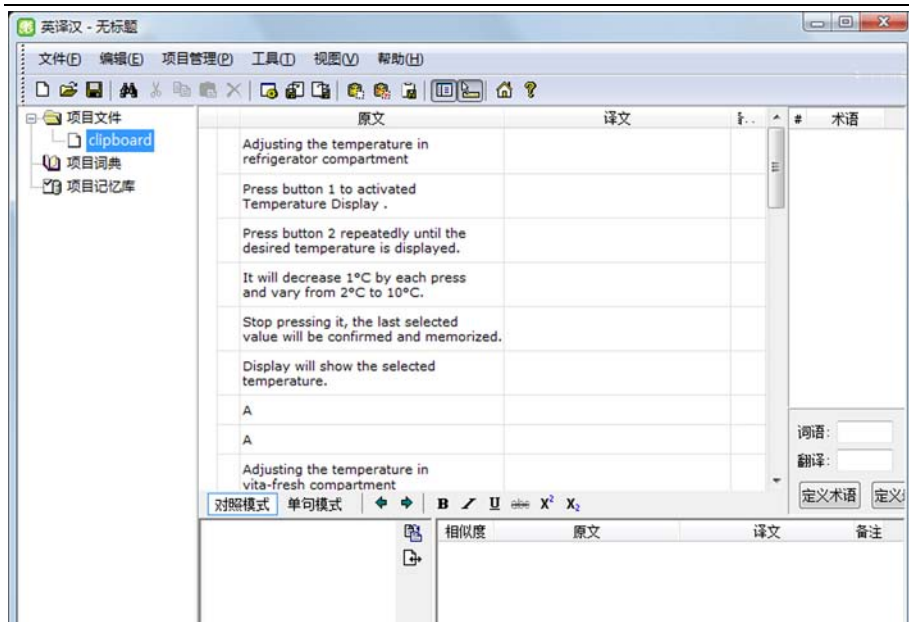
#### Adjusting the temperature in vita-fresh compartment

- Press button 1 to activate Temperature Display (1).
- Press button 2 repeatedly until the desired temperature is displayed. It will decrease 1°C by each press and vary from 4°C to 0°C.
- Stop pressing it, the last selected value will be confirmed and memorized. Display (1) will show the selected temperature of Bio-fresh compartment.



#### Adjusting the temperature in freezer compartment

- Press button 1 to activate Temperature Display (1).
- Press button 2 repeatedly until the desired temperature is displayed. It will decrease 1°C by each press and vary from -18°C to -26°C.
- Stop pressing it, the last selected value will be confirmed and memorized. Display (1) will show the selected temperature of freezer compartment.



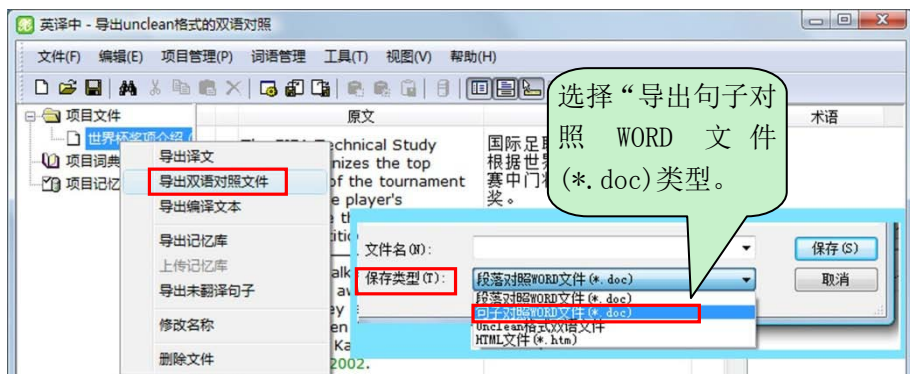
有了剪贴板的支持，任何可以编辑的文件格式（即任何可以复制、粘贴出来的内容）都可以利用雪人来翻译，为了让更多的译员体会到剪切板带来的方便，我们将这个功能同时对免费版的用户开放。



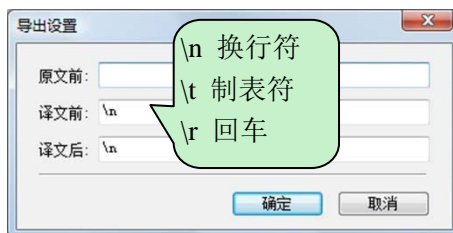
## 第五节 导出句子双语对照和 Unclean 格式文件

### 一、导出 WORD 格式的句子双语对照

雪人在 V1.27 版中增加了导出句子格式的双语对照格式，双语对照以句子为单元，导出的文件格式支持：TXT、WORD、EXCEL、PPT 格式。右键点击左侧“项目文件”下面的翻译文档，选择【右键菜单→导出双语对照文件】功能，如下图：在弹出的“导出双语对照文件”对话框中，点击“保存类型”的下拉框，选择“导出句子对照 WORD 文件(\*.doc)”。

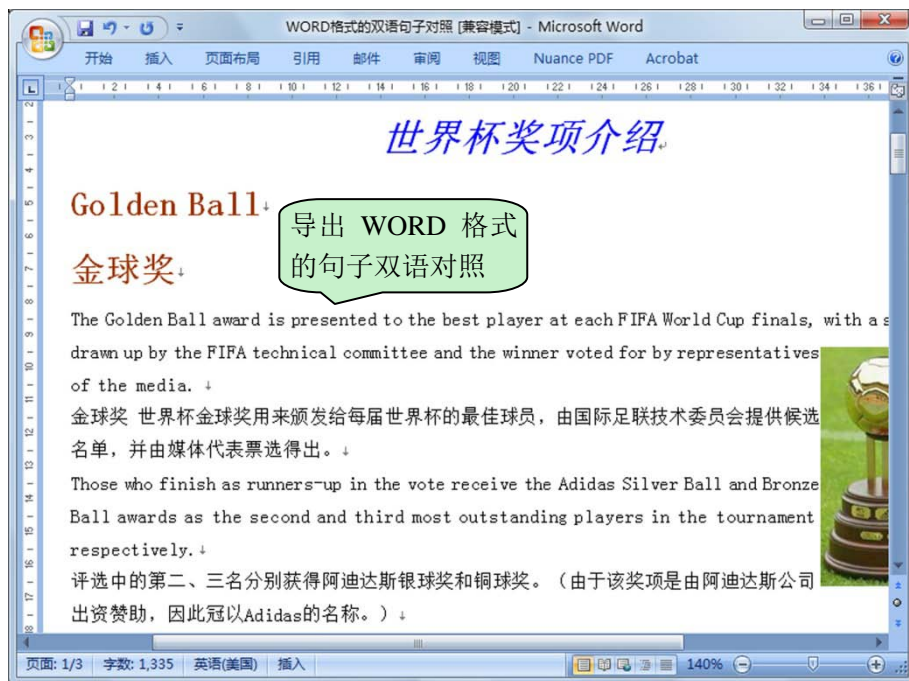


点击“保存”按钮后，会弹出“导出设置”窗口，如下图：用户可以自己定义原文、译文前后的分隔符号，如下图：\n 插入换行符，\t 插入制表符（Tab），\r 插入回车。若导出文本格式时，通常需要写成\r\n。还可以输入任意文字，例如可以通过添加括号，将译文部分括起来。



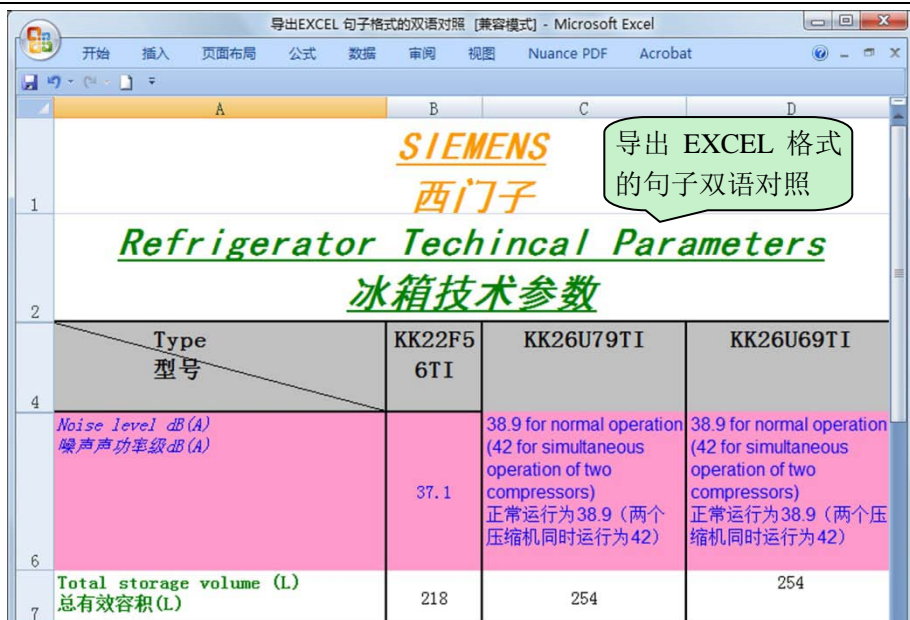
导出的 WORD 格式的句子双语对照如下图：对于 WORD 中插入的表格，如果希望导出段落格式的双语对照，可以将雪人的 STP 项目文件另存一份后（避免修改原文件），利用合并某些句子的方法来实现。





## 二、导出 EXCEL 格式的句子双语对照

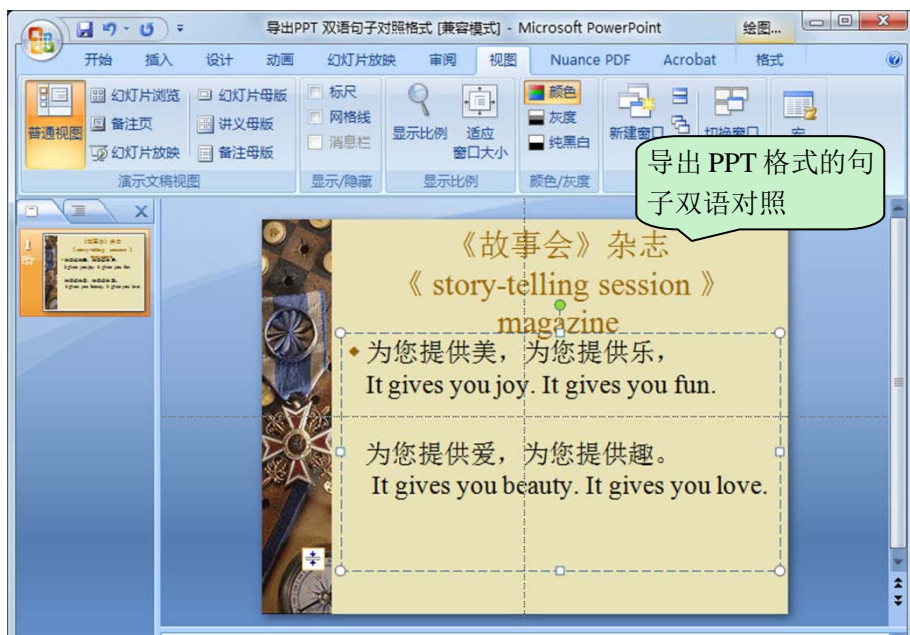
如果是 EXCEL 文件，导出的句子双语对照文件格式如下图：



Type 型号	KK22F5 6TI	KK26U79TI	KK26U69TI
Noise level dB (A) 噪声功率级dB (A)	37.1	38.9 for normal operation (42 for simultaneous operation of two compressors) 正常运行38.9 (两个 压缩机同时运行42)	38.9 for normal operation (42 for simultaneous operation of two compressors) 正常运行38.9 (两个 压缩机同时运行42)
Total storage volume (L) 总有效容积 (L)	218	254	254

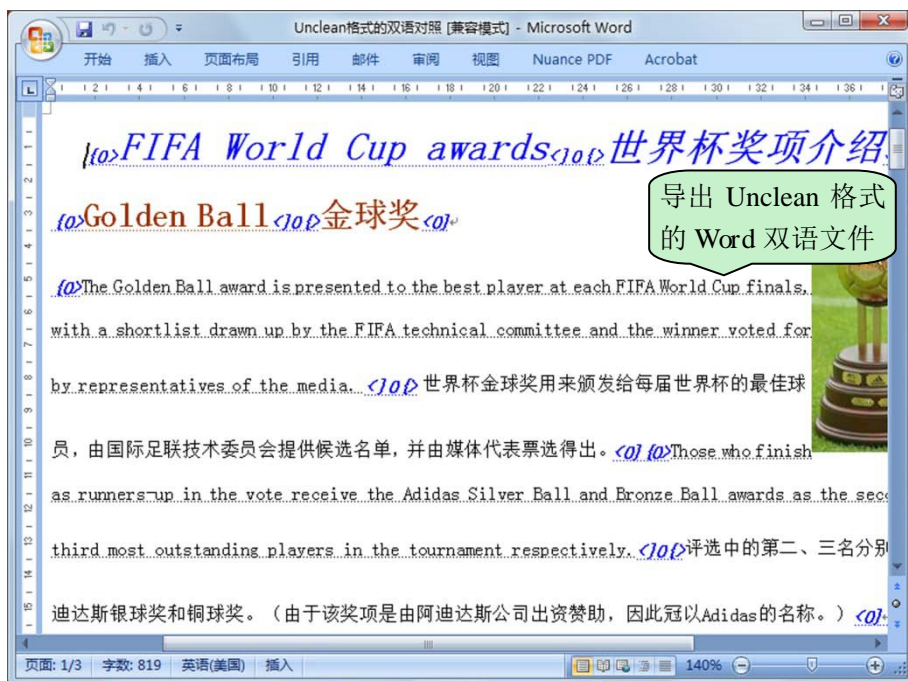
### 三、导出 PPT 格式的句子双语对照

如果是 PPT 文件，导出的句子双语对照文件格式如下图：



## 四、导出 Unclean 格式的 Word 双语文件

还可导出 Unclean 格式的双语文件，方便使用雪人软件的用户与使用其他 CAT 软件的用户之间交流。右键点击左侧“项目文件”下面的翻译文档，选择【右键菜单→导出双语对照文件】功能，如下图：在弹出的“导出双语对照文件”对话框中，点击“保存类型”的下拉框，选择“Unclean 格式的双语文件”即可。导出的结果如下图：

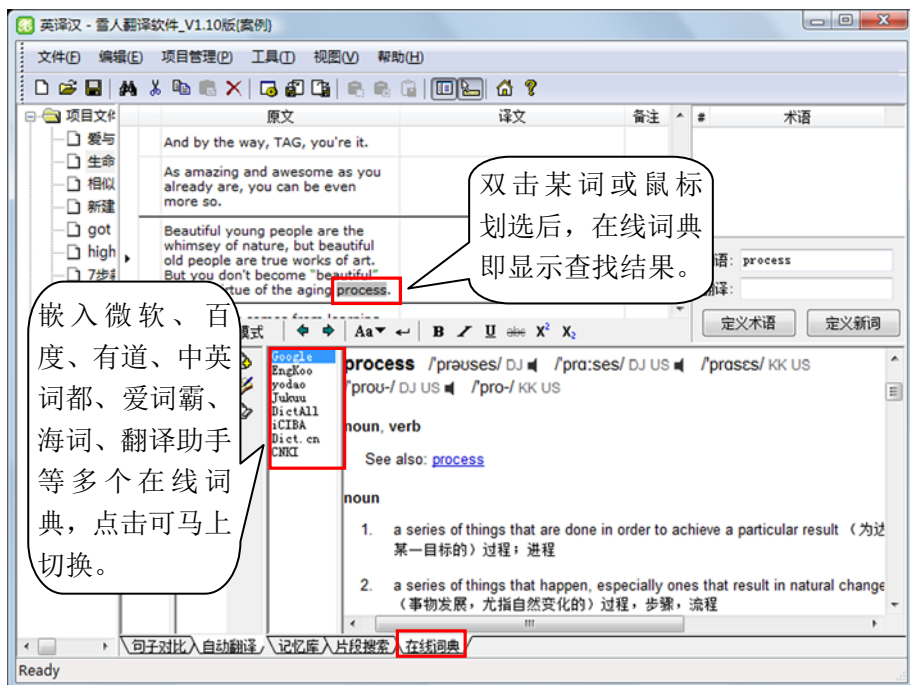


# 第五章 网络资源的整合运用

## 第一节 嵌入多个在线词典

雪人系统自带词典非常丰富，标准版有 1200 万、涵盖 60 多个专业领域的词库（免费版有 30 万的词汇）。同时，还很好地利用了在线词典的资源为译员服务，雪人在软件中嵌入“[EngKoo](#)、[Baidu](#)、[youdao](#)、[Jukuu](#)、[iCIBA](#)、[Dict. cn](#)、[DictAll](#)、[CNKI](#)、[FastDict](#)、[CiBO. cn](#)”等 10 个在线词典，译员只要在原文中[双击或用鼠标划选](#)某词语后，立即显示查找的结果，点击即可切换到其他在线词典查看。

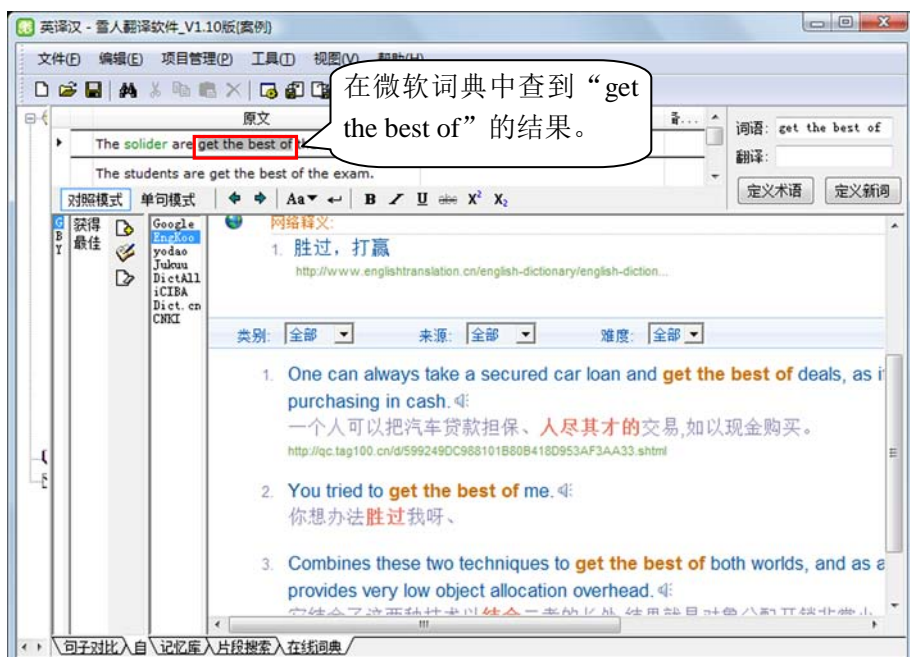
下面我们来看看具体的操作：点击屏幕下方的[在线词典](#)页签，用鼠标双击原文的某单词或用鼠标划选，例如：选择了“President”，在“Google 在线词典”中就立即显示出 President 的查找结果，软件默认显示的“Google 的在线词典”，第二是“微软的 EngKoo”、第三是“有道词典 youdao”、第四是“中英 Jukuu”、第五是“词都 DictAll”、第六是“爱词霸 iCIBA”、第七是“海词 Dict. cn”、第八是“翻译助手 CNKI”，点击它们可立即切换到该词典并同时显示查找结果。



接下来我们来查找一下：“get the best of”这个短语该如何翻译，用鼠标划选之后，在 Google 的词典中查找不到相应的解释，如下图：



如果对在“Google 的词典”中查找的结果不满意，可以立即切换到其他的在线词典中查找，例如切换到“微软的 EngKoo”中看看它对“get the best of”这个短语的解释，此时只要点击左侧的“EngKoo”即可。如下图：

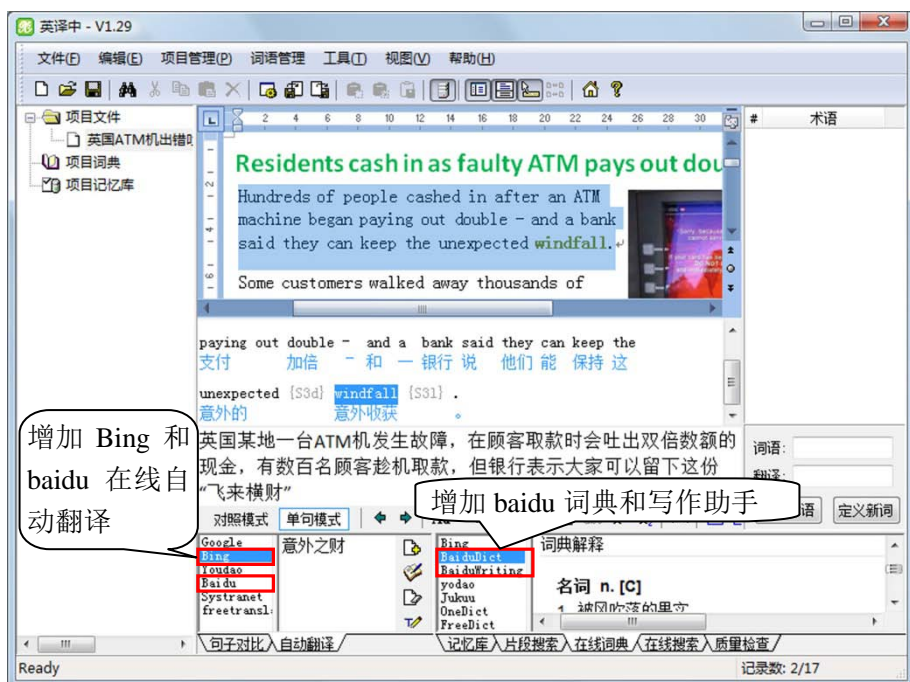


Google 和微软的“EngKoo”另一个有用的功能是可以对短语、短句进行机器翻译，例如上面当我们在原文处划选“get the best of”时，“EngKoo”给出的是这个短语的例句参考，如果我们在原文中选择整句“The solidier are get the best of the war.”，如下图，这一句的机器自动翻译结果是“士兵都得到最好的战争”。

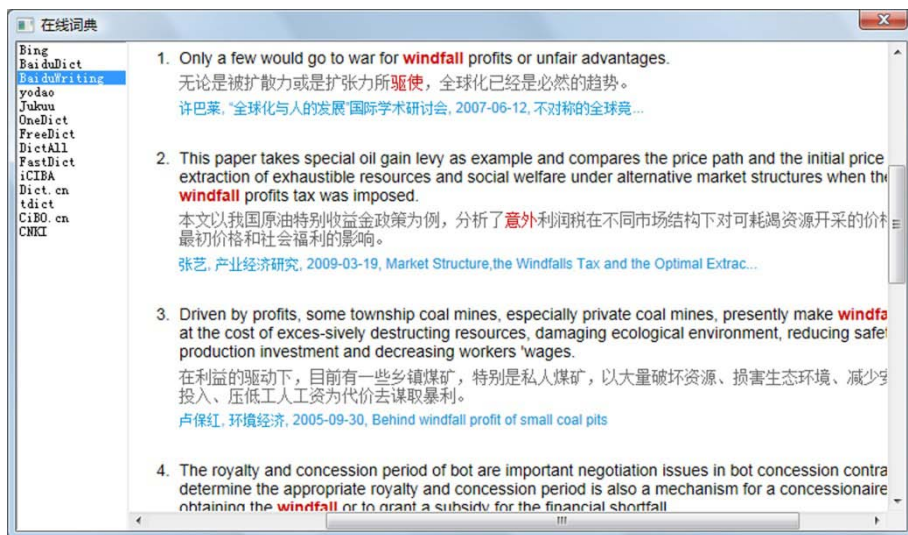




另外，在 V1.29 版中，增加了“Baidu 词典 (BaibuDict)”和“Baidu 英语论文写作助手 (BaiduWriting)”连个在线词典。标准版还增加了“Bing 和 Baidu”两个在线自动翻译。用户可以根据需要选用。



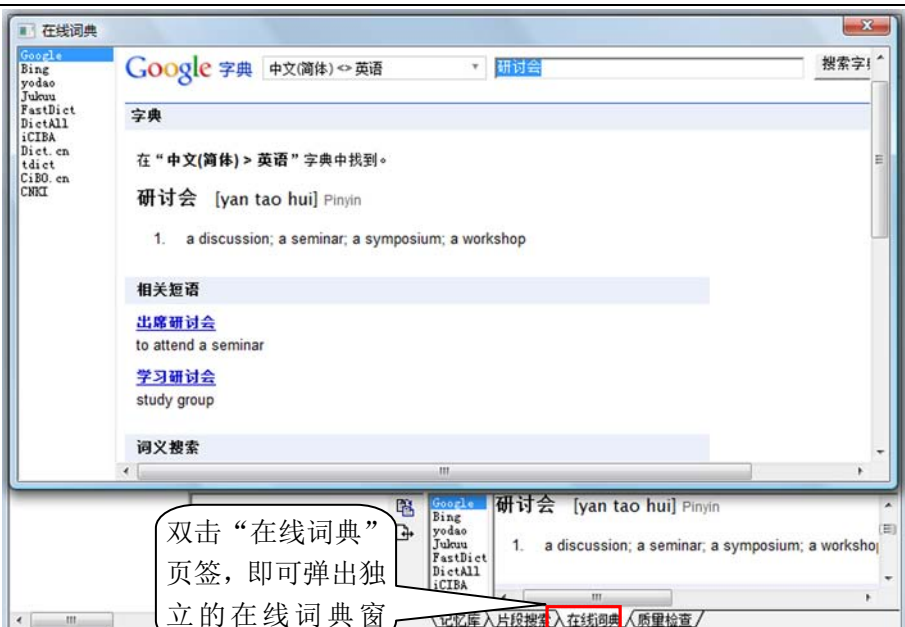
“Baidu 英语论文写作助手(BaiduWriting)”中有不少的例句可以查询，会给翻译工作带来一定的参考作用。如下图：



## 弹出独立的在线词典窗口

雪人的鼠标轻轻一划马上查词的“在线词典”功能深受译者的喜爱，但是由于屏幕的空间有限，屏幕右下角的“在线词典”窗口不够大，经常需要来回滚动。为解决这个问题，雪人在 V1.18 版中增加了“弹出独立的在线词典”窗口（**双击**屏幕下方的“**在线词典**”页签，即可弹出该窗口），这样使用户查词变得更加方便。如下图：





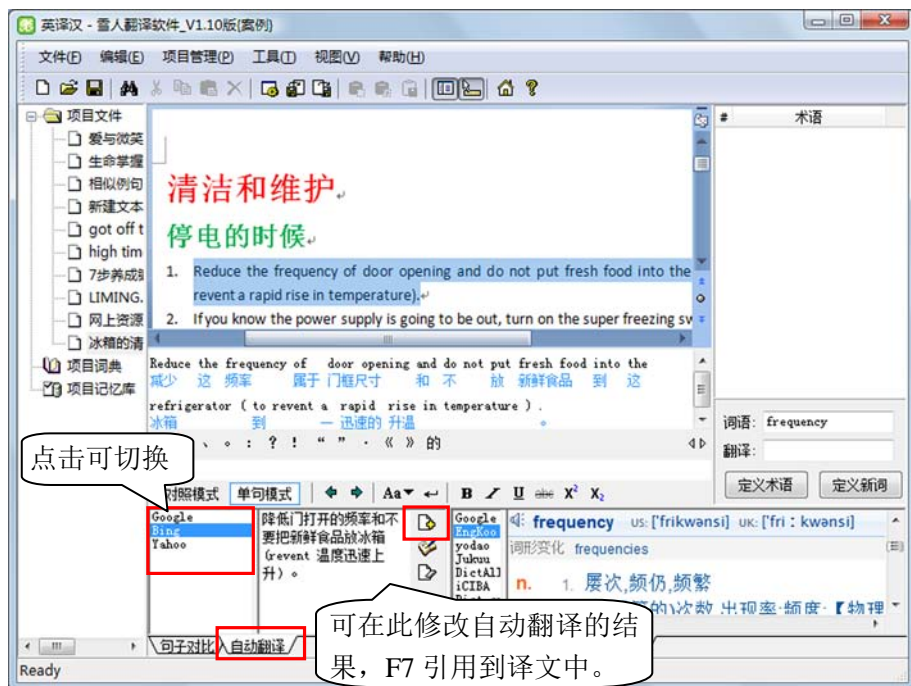
## 第二节 嵌入在线自动翻译

或许有人会认为机器翻译——那是给不懂英语的人看的，对专业译员根本无用。虽然机器自动翻译的准确性目前还不尽人意，但这并不意味着机器自动翻译对译员来说无用，它可以在一定程度上给我们提供参考及减少打字的工作量。如果希望直接使用机器自动翻译的结果作为译文，其准确性显然难以达到要求；但如果在其基础进行修改，会比直接录入要快很多。尤其是目前几个主流的网上在线翻译对简单的短句翻译效果还是不错的，而且它们还存在有巨量的句库，如果它们句库中存在的句子，通常会翻译的比较好。我们应该好好地利用它们来提高我们的工作效率！

如何高效地利用这些免费的资源呢？如果按传统的方式：即复制、粘帖一句到某个在

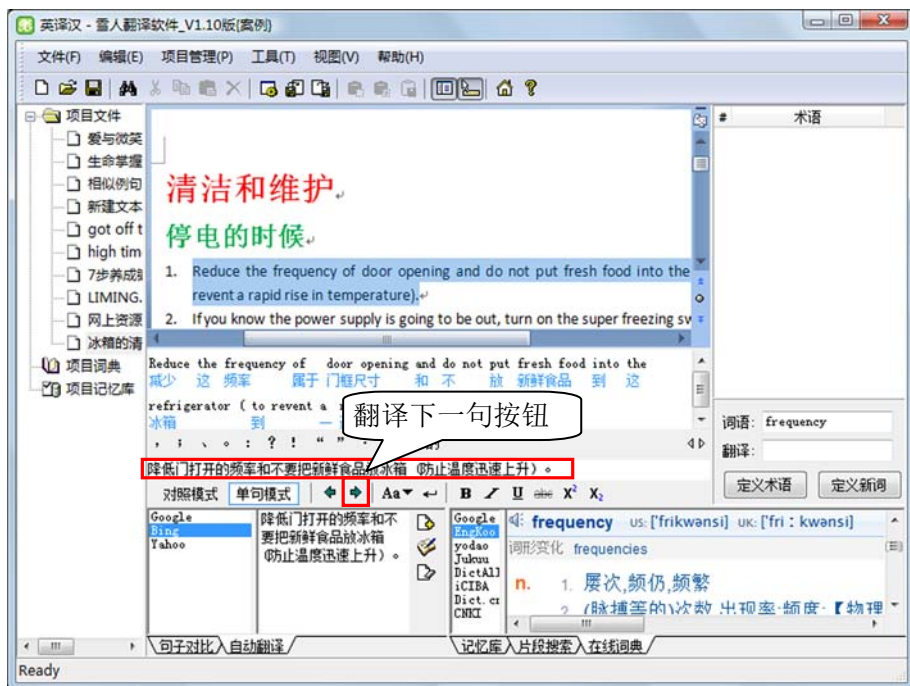
线翻译的网站，点击它的翻译按钮进行翻译，若对其翻译结果满意就再次复制、粘贴回来；如果对其翻译结果不满意，则再去其他在线翻译的网站看看，但有时可能还存在各个在线翻译的结果都不好，如此这般工作则全部白费了。这对时间紧、任务重的译员来说，简直是在浪费时间，还不如自己直接录入翻译。然而，雪人翻译软件却提供了一个高效的工作方法，不同麻烦它既可以很好地利用在线翻译，又可以马上对在线翻译的结果进行修改；既可以整句的翻译，又可以分片段进行翻译；最重要的是利用这些在线翻译时，并没有增加多少额外的工作量。下面我们将通过实例来讲解雪人是如何利用在线翻译的？


点击屏幕下方的**自动翻译**页签，当翻译到某一句时，“自动翻译”窗口就自动给出了“在线翻译”的结果，雪人目前支持**Google、Bing、Yahoo、Youdao、Systrant、freetranslation**多个在线自动翻译引擎，如果对某个引擎的翻译结果不满意，可立即切换到其它的引擎中。如下图：




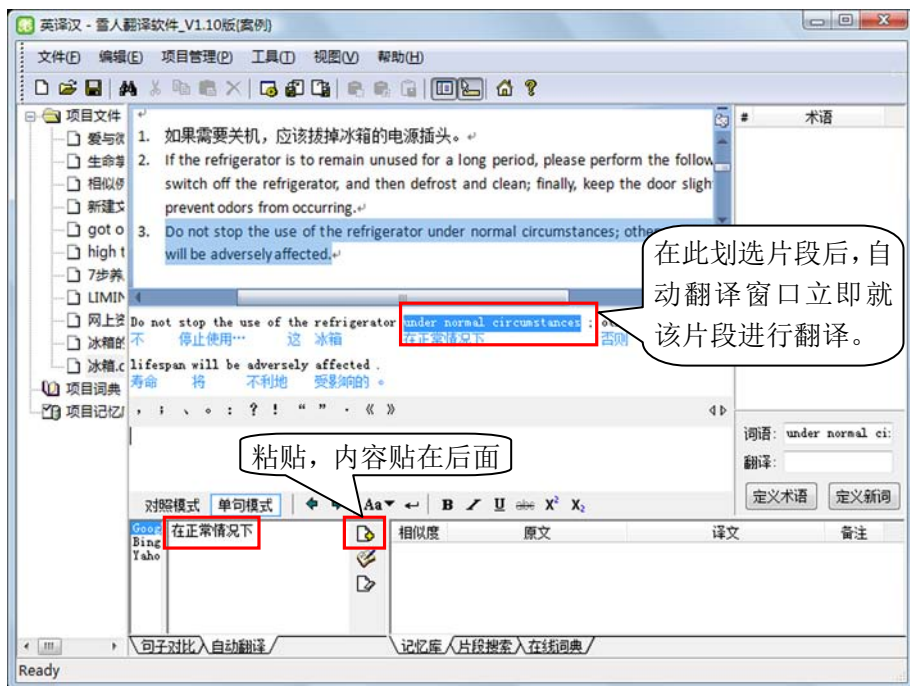
译员可以直接修改自动翻译的结果，修改好后点击旁边的**粘贴**按钮或按【F7】键将其引用到上方的译文区，例如：对于“Reduce the frequency of door opening and do not put fresh food into the refrigerator (to prevent a rapid rise in temperature).”这一句，自动翻译结果为：“降低门打开的频率和不要把新鲜食品放冰箱 (prevent 温度迅速上升)。”自动翻译的结果基本可以，我们略作修改后就按【F7】将它引用到译文中了。

如下图：



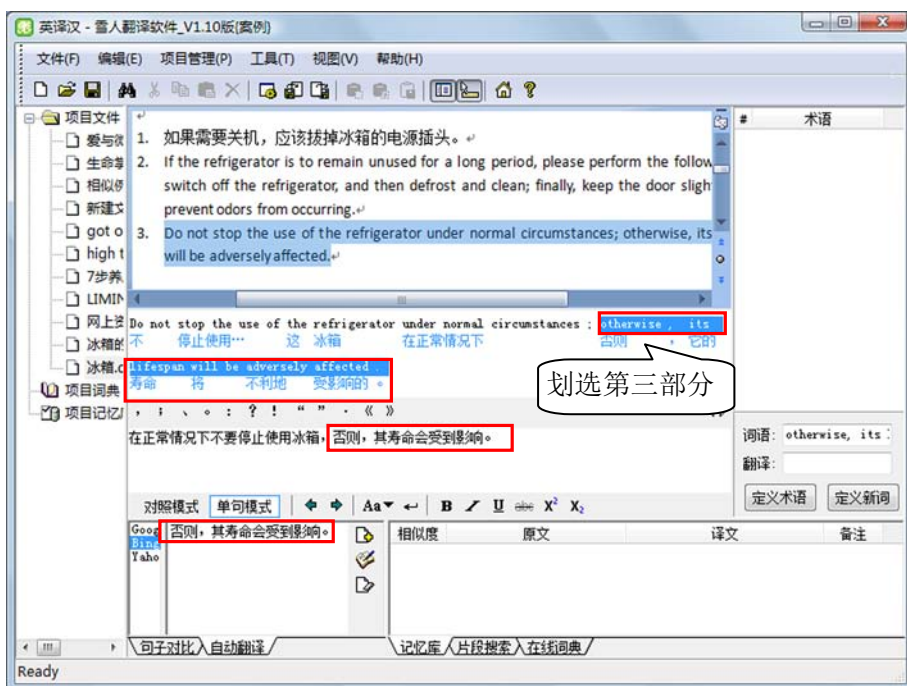
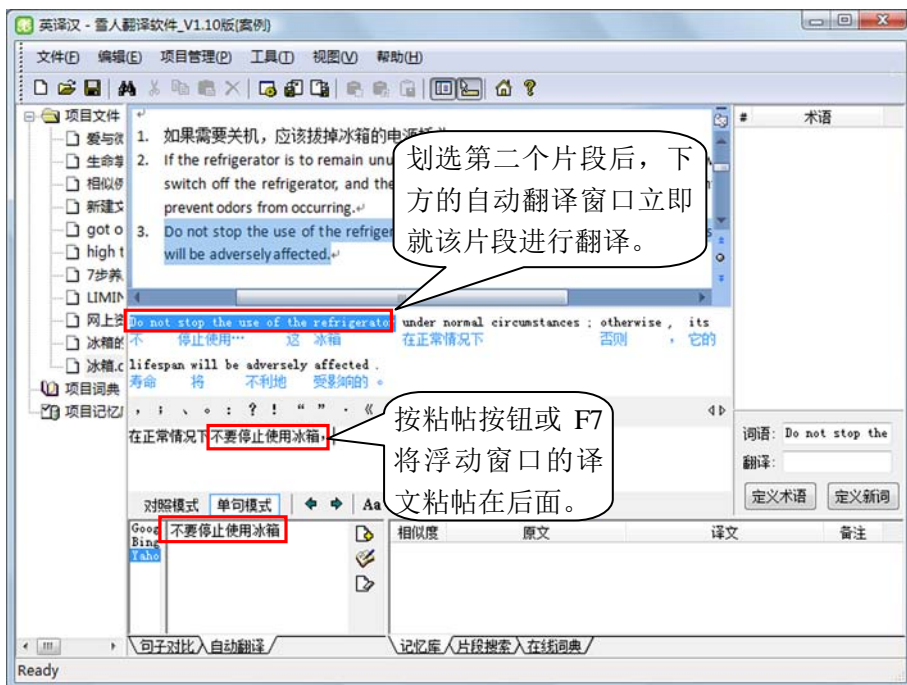
翻译好上一句后，点击“ 下一句”按钮就可以直接控制原文的下一句，此时在自动翻译窗口中立即出现下一句的自动翻译结果。

自动翻译对长句的翻译结果往往不是很理想，但若将它拆分成几个短句或片段时，其准确性还是不错的。例如：当翻译到“Do not stop the use of the refrigerator under normal circumstances; otherwise, its lifespan will be adversely affected.”这一句时，整句的自动翻译结果并不是很通顺。此时若我们将它拆分成几个片段来看看它的翻译结果，首先从语序来说我们知道应该先翻译“under normal circumstances”这个片段，所以我们先用鼠标在下方划选了该片段，划选后在自动翻译窗口中立即将此片段翻译成了“在正常情况下”，如下图：翻译结果不错，点击 粘贴按钮或按【F7】键将它复制到上方的译文区中。



接下来的语序应该是：“Do not stop the use of the refrigerator”，所以我们在下方划选了该片段，划选后在浮动窗口中立即将此片段翻译成了“不要停止使用冰箱”，如下图：翻译结果不错，我们点击浮动窗口的 **粘贴** 按钮或按【F7】键将它复制到下方的译文区中，**粘贴** 按钮是将浮动窗口的内容紧贴在原来的后面，而 **清除** 按钮是将翻译的该句清除，译员需要按正确的语序重新翻译，而按【F7】键软件会自动判断应该整句覆盖还是粘贴在其后面。

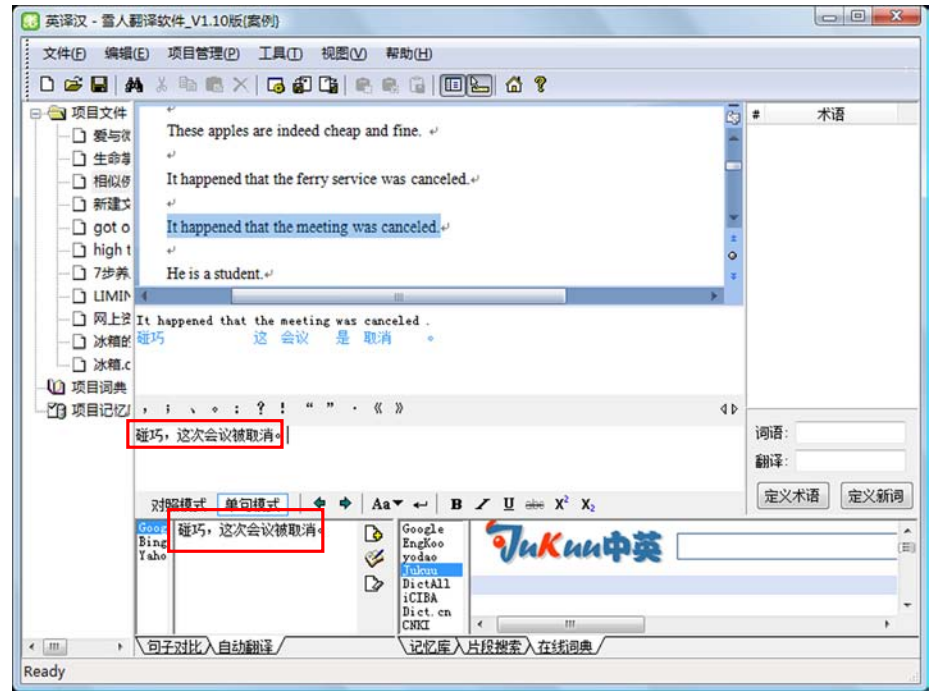
第三部分划选“otherwise, its lifespan will be adversely affected.”划选后在浮动窗口中立即将此片段翻译成了“否则，其寿命会受到影响”，如下图：翻译结果不错，我们点击 **粘贴** 按钮或按【F7】键将它粘贴到后面。此时，整句翻译全部完成了，全部译文来自于自动翻译，我们没有手工录入一个文字，只是将以整句拆分成了3个片段，划选了3次即完成了该句的翻译工作，这显然比完全手工录入要快的多。所以说，利用在线自动翻译是可以很好地提高工作效率的，关键是使用的方法及工具，希望译员可以灵活运用好这个资源。








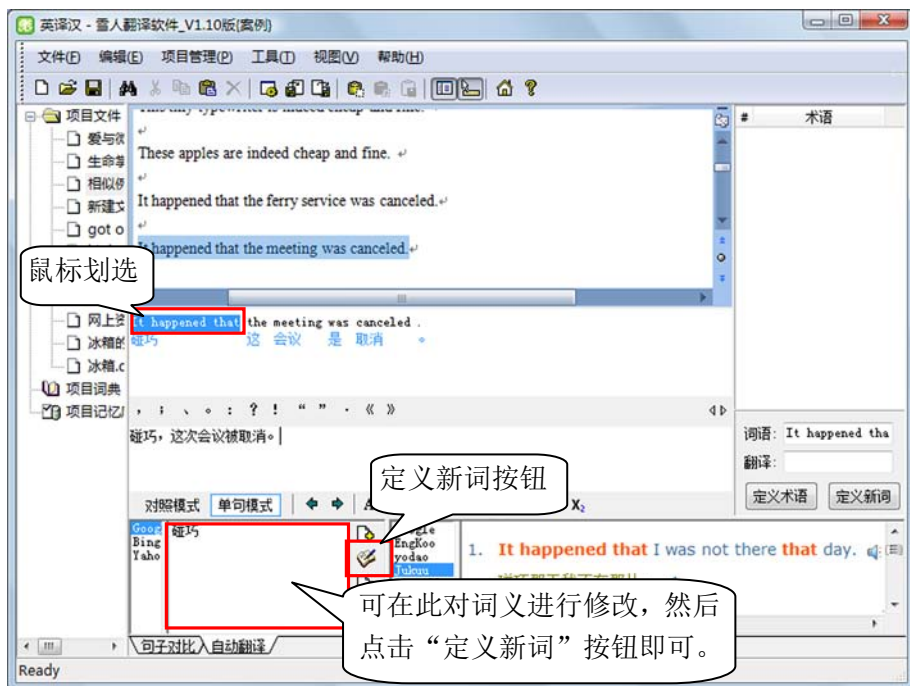
## 加入到词典

为了提高翻译的工作效率，翻译时应该注意短语、常用片段的积累。利用在线自动翻译可以很方便地实现这个积累，例如当翻译 “It happened that the meeting was canceled.” 这一句时，自动翻译窗口给出了 “碰巧，这次会议被取消。” 的不错译文：如下图所示：



我们按  粘贴按钮或【F7】键引用下来后，觉得 “It happened that” 是个常用的句型，希望将它保存起来，对此，雪人提供非常方便的功能，如下图；具体操作如下：

1. 用鼠标在中部的 “词语自动翻译区” 中划选 “It happened that”
2. 此时，在 “自动翻译窗口” 立即出现其译文 “碰巧”，若认为不确切，还可以在此对译文进行修改。
3. 点击旁边的  定义新词按钮，即可将它加入到  项目词典。



下图显示出了刚才添加到项目词典中的词语。

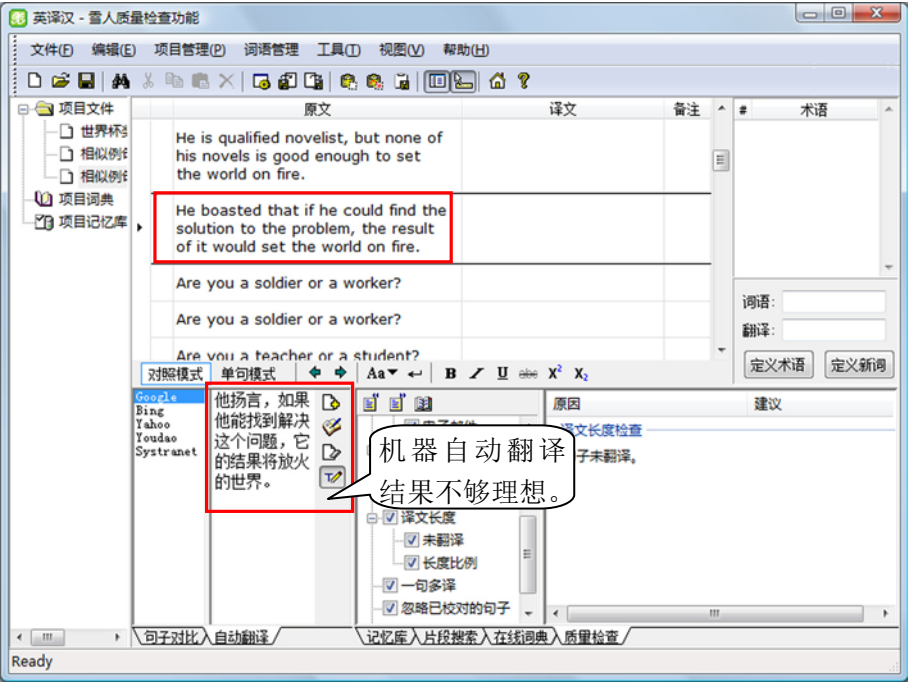
序号	原文	词性	译文	备注
1	got off to a flying	开门红		
2	pan out	一帆风顺		
3	president	总经理		
4	It happened that	碰巧		



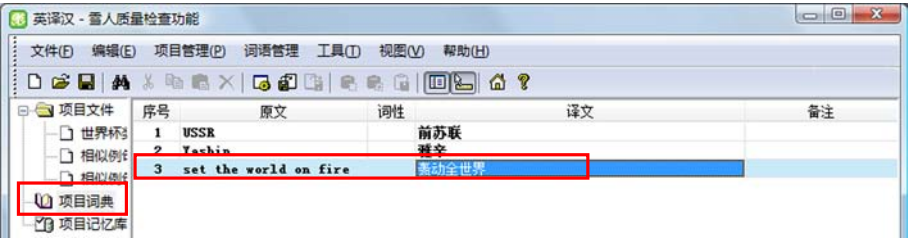
### 第三节 用户术语库与在线自动翻译的结合


支持使用本地术语库与在线翻译混合的翻译模式，通过使用自定义的术语，使得在线自动翻译更加准确，进一步提升自动翻译的质量。

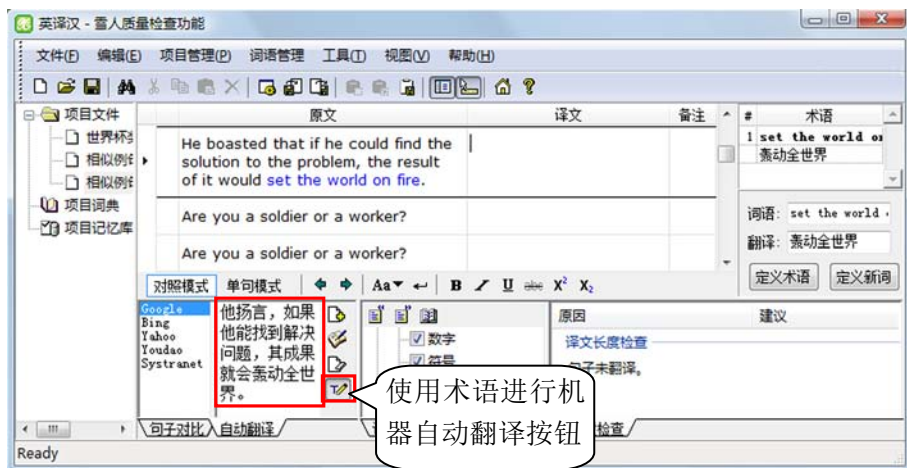
例如 “He boasted that if he could find the solution to the problem, the result of it would set the world on fire.” Google 在线自动翻译给出的译文是：“他扬言，如果他能找到解决这个问题，它的结果将放火的世界。” 显然不够通顺，原因是 “set the world on fire” 机器未能对它进行正确的翻译，如下图：



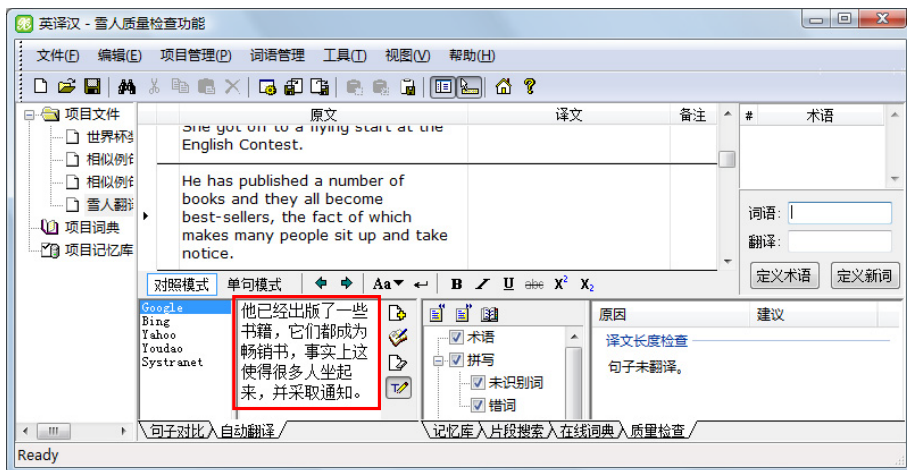
此时，若我们将 “set the world on fire 轰动全世界” 这个词语加入到 “项目词典”，并将它设为 “术语”（术语以粗体显示，以区别于普通词汇）。如下图：



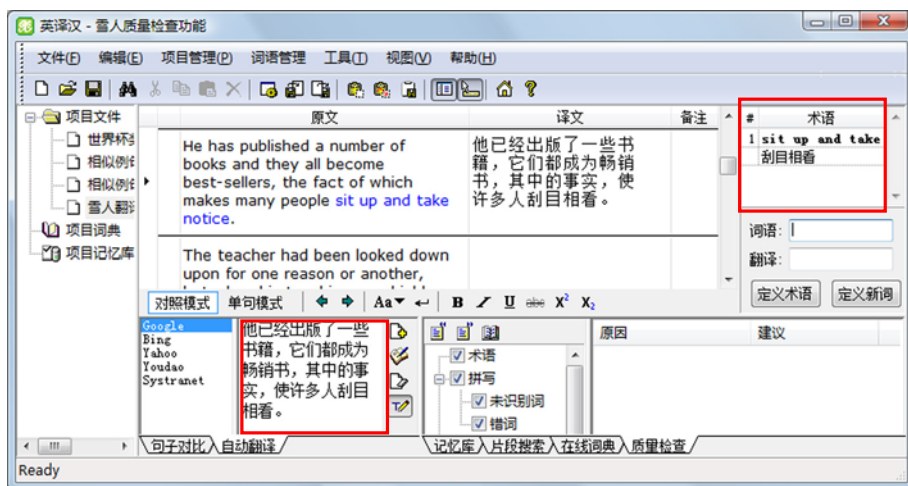
然后，我们点击自动翻译窗口旁边的“使用术语翻译”的按钮（该按钮鼠标点击后会稍微凹下去，以后各句均表示使用术语进行翻译，再点击则表示不使用用户术语库）。使用术语进行翻译，此时 Google 给出的自动翻译结果修正为“他扬言，如果他能找到解决问题，其成果就会轰动全世界。”这显然要好很多了。



再如这一句“He has published a number of books and they all become best-sellers, the fact of which makes many people sit up and take notice.” Google 的自动翻译译文为“他已经出版了一些书籍，它们都成为畅销书，事实上这使得很多人坐起来，并采取通知。”如下图：



但当我们把“sit up and take notice 刮目相看”添加到术语库后，此时 Google 的自动翻译译文为“他已经出版了一些书籍，它们都成为畅销书，其中的事实，使许多人刮目相看。”



由此可见，利用自己定义的术语库结合 Google 等的在线自动翻译，将可以很好地改善机器自动翻译不能按自己的要求正确取词义的问题，使得自动翻译的质量更上一筹！达到减少文字的录入速度。（自动翻译仅对标准版的用户提供，免费版的用户没有这个功能）。

## 第四节 在线搜索

雪人的“在线词典”功能因其使用的方便性而深受译员们喜爱；但是对于有些词汇（例如：人名、地名、机构名称等等），有时“在线词典”亦找不到相应的结果；此外，对于有些词汇，仅知道其词义可能远远不够，我们希望了解与之相关的更多知识，例如：人物的背景资料、行业的规范标准等等。这时就需要登录 Google、Wikipedia（维基百科）、百度、Hudong（互动百科）等网站，在它们海量的信息库中去搜索自己想要的结果了。

雪人在 V1.22 增加了“在线搜索”功能，将 Google、Wikipedia（维基百科）、百度、Hudong（互动百科）、csres（工标网）等一些翻译人员常用的工具网站嵌入到雪人 CAT 中，鼠标在原文划选词语后，不但能从“在线词典”中查询结果，点击屏幕下方的“在线搜索”页签，即可显示从“Google、Wikipedia”等网站搜索到的相关结果而无需再次划选或输入，鼠标一次划选、多次搜索，十分方便。如下图：



在屏幕下方的“在线词典”旁边增加了新的页签——“在线搜索”，例如上图中：鼠标划选了“Prince Willian”这个词语后，“在线搜索”窗口立即显示“Wikipedia”关于“Prince Willian”的介绍文章，如果觉得下方的屏幕空间不够大，可以双击“在线搜索”页签，弹出独立的放大窗口来显示搜索到的内容，如下图：



点击左侧的“Google”，可显示“Prince William”从“Google”上搜索到的内容。如下图：另外，如果你发现其他对翻译搜索有帮助的网站，欢迎向我们反映，我们会考虑将它添加到软件中。





# 第六章 预翻译

利用记忆库、词典、规则词典、本地术语库和在线翻译相结合等方式进行预翻译，快速生成译文。

## 一、利用记忆库及规则词典预翻译

雪人在 V1.21 版中增加了对“预翻译”的支持，可以使用记忆库、词典、规则词典及在线自动翻译对文章进行快速的预处理，如果记忆库中存在已翻译的句子，预翻译时则优先选用记忆库中的句子，并对相应的译文来源进行标注，方便用户对它们进行校对。利用记忆库和规则词典相结合的预翻译方式，可以对一些内容相差不大的文档和格式化的文档进行快速的翻译。

例如下面两篇关于“展会邀请”格式文本的翻译，内容基本相同，不同之处是展会的日期，举办的城市、展会的期数、展会大厅、展台等等（见示例文件“预翻译案例”）。

布朗先生：

如您所知，在 2011 年 6 月 1 日至 2011 年 6 月 18 日期间将在广州举办第 110 届广交会。

您肯定已有一份参观广交会的时间安排了。我们非常期待您能安排参观位于第 6 号大厅，展位号为 18 的展台。

最好今天就能回复随函附送的邀请函。

我们期待您的前来参观。

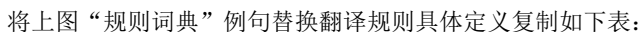
安妮谨上

对于各种展会的名称、各个大中城市等我们可以预先在项目词典中输入，做成展会词典库，如下图：

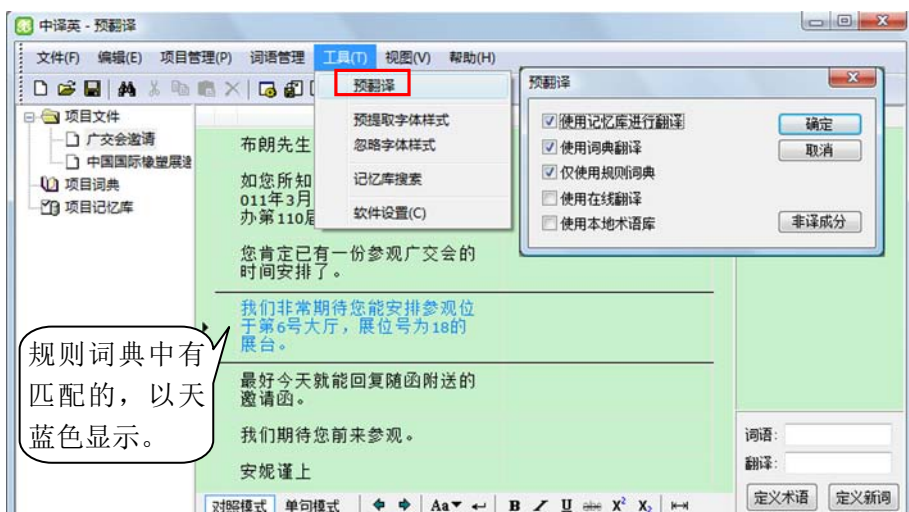


对于中英日期的翻译，我们可以在规则词典中定义日期的规则，就可以实现所有日期的自动翻译。并将例文中一些需要替换的地方也在规则词典中定义（关于规则词典的详细

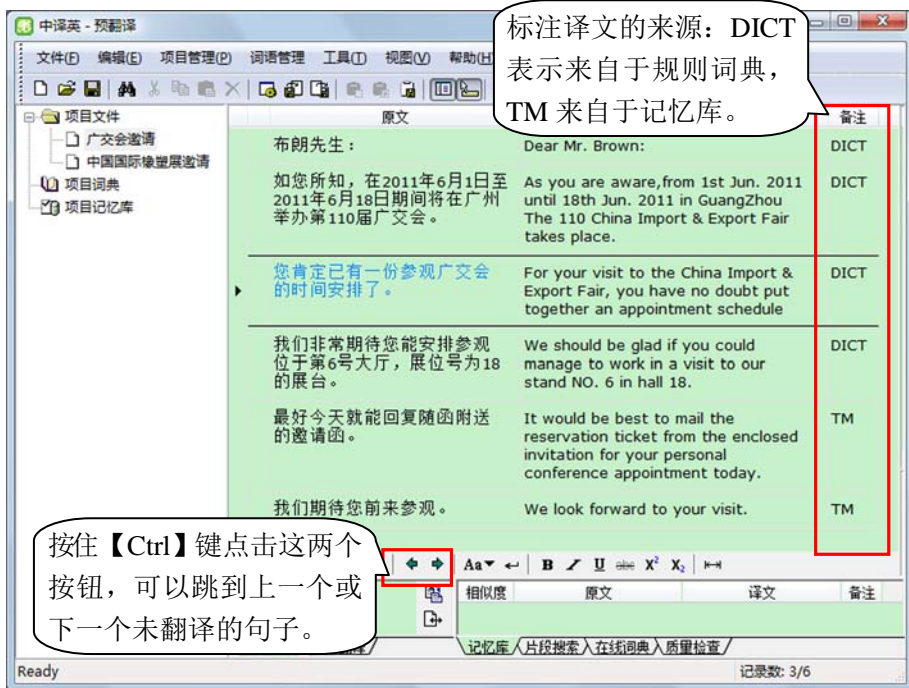




定义好规则后，我们看到，如果是规则词典中有匹配的，会以天蓝色显示，点击【工具菜单→预翻译】，弹出“预翻译”窗口，如下图：



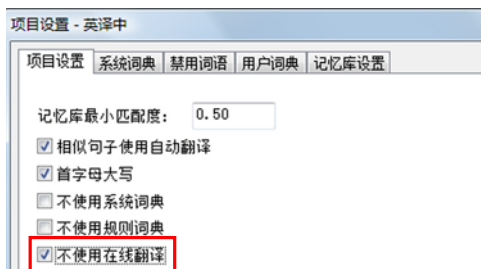
上图是利用记忆库、词典、规则词典进行预翻译（使用在线翻译及本地术语库这两选项没有勾选，表示不使用这两个功能），点击“确定”按钮后，瞬间生成了高质量的译文，如下图：软件对预翻译的每一句译文都做了标注，方便用户鉴别，“DICT”表示来自于规则词典，“TM”来自于记忆库、“GOOGLE”表示来自于 GOOGLE 的在线自动翻译。



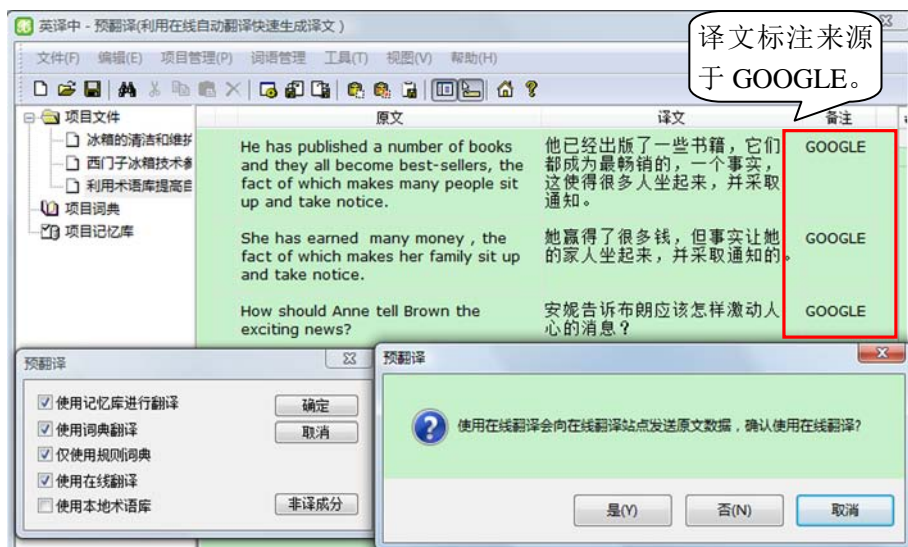
## 二、利用在线翻译进行预翻译

除了利用记忆库、词典、规则词典进行预翻译外，还可以利用“在线自动”翻译进行预翻译。默认状态下“在线翻译”的选项并没有打开，在“预翻译”窗口中勾选“使用在线翻译”这个选项可以启用这个功能，由于使用在线翻译需要向在线翻译网站发送原文，此时会弹出一个提示窗口提醒用户，如果是保密文档，则不建议使用这个功能。

保密说明：软件除了在使用“在线词典和在线翻译”时才向互联网发送请求翻译的内容，除此之外，不会向互联网发送任何数据。为了避免在线翻译意外发送了待翻译的句子，雪人CAT在V1.21版后，新建的项目文件默认不启用“在线翻译”功能，如果用户需要使用该功能，可以点击【项目管理→项目设置】设置为使用在线翻译。如右图：

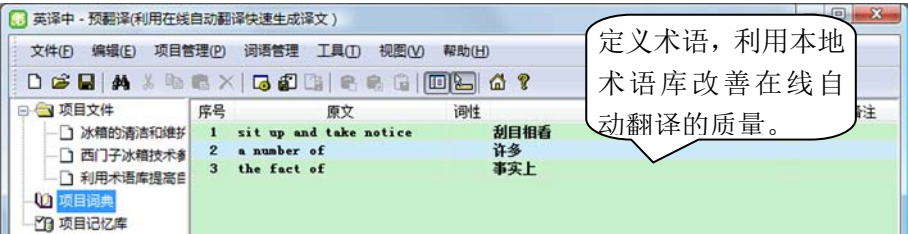


确认后软件将利用 GOOGLE 的在线翻译，生成译文，并在备注栏中注明译文的来源是“GOOGLE”。如下图：

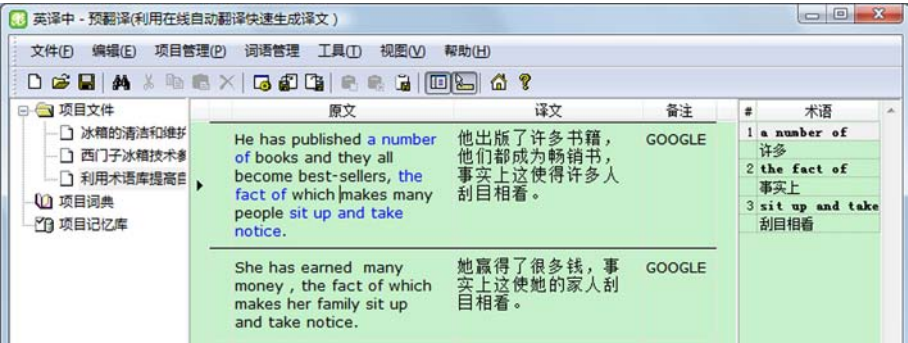


如上图所示，自动翻译的译文并不理想，例如：“He has published **a number of** books and they all become best-sellers, **the fact of** which makes many people **sit up and take notice.**” 这一句，GOOGLE 自动翻译的译文是：“他已经出版了一些书籍，它们都成为最畅销的，一个事实，这使得很多人坐起来，并采取通知。”我们发现译文不好的原因是

GOOGLE 对 **a number of**、**the fact of**、**sit up and take notice** 这几个短语的取词不正确，我们将这几个词作为术语定义在雪人 CAT 的“项目词典”中，如下图：

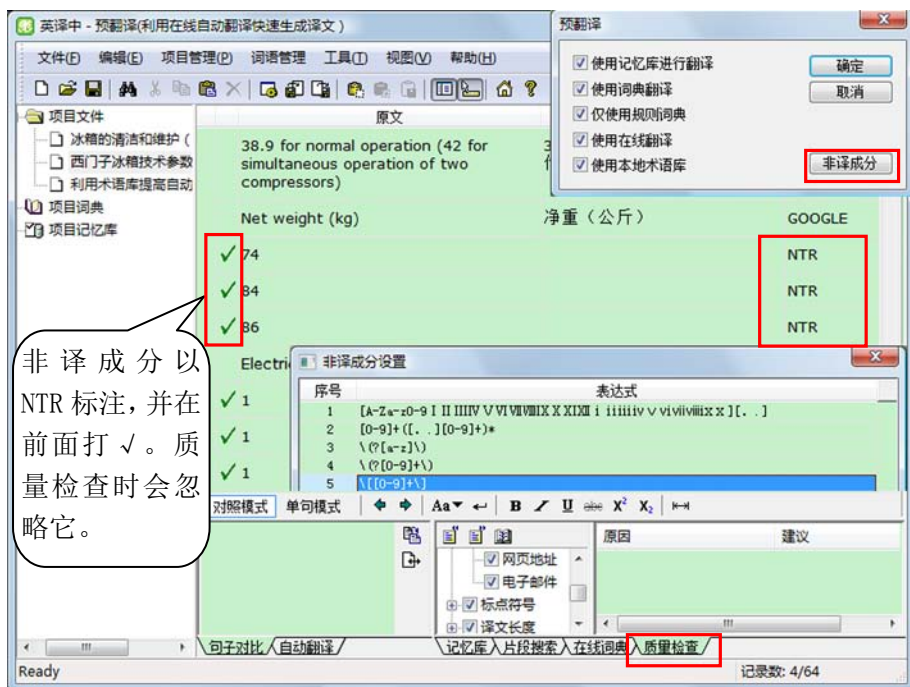


然后将原来的译文删除（若已经有译文的，预翻译则不再填写译文），重新进行预翻译，并勾选“预翻译”窗口的“使用本地术语库”这个选项，我们发现，利用本地术语库可以很好地改善 GOOGL 自动翻译的质量，现在，上句的译文变成了“他出版了许多书籍，他们都成为畅销书，事实上这使得许多人刮目相看。”如下图：



### 非译成分

在预翻译窗口中，有一个“非译成分”按钮，用于定义不需要翻译的内容，可以使用正则表达式来定义匹配规则。定义“非译成分”后，预翻译时软件会在备注栏以“NTR”标注，并在该句前面打“√”，表示该句的译文完全正确。软件将认为该句使用原文进行翻译，质量检查时不会被认为是未翻译的句子。见下图：



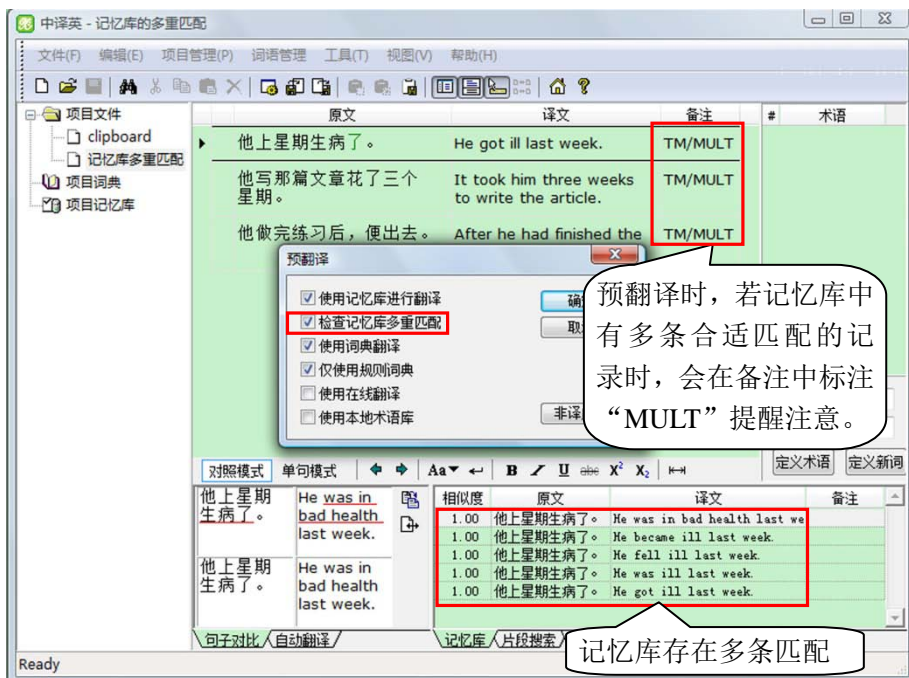
**注意：**预翻译的结果不会直接加入到项目记忆库中。预翻译不处理原文中的样式代码，如果原文中含有样式代码，需要使用质量检查工具检查是否有遗漏；Google 的在线翻译不能识别纵向制表符（VT），包含纵向制表符的句子会翻译为乱码。

利用记忆库、词典、规则词典、本地术语库和在线翻译向结合等方式进行预翻译，对于一些格式文档，可以快速生成译文。预翻译也可以用于处理重新翻译修改后的文档。重新导入修改后的文档，使用预翻译可以自动填充之前已翻译好的内容。希望用户能触类旁通、灵活运用，创建合适自己的翻译模式，更好地提高翻译效率！



### 三、检查记忆库多重匹配

新版本的预翻译中加入了“检查记忆库多重匹配”选项，用于检查记忆库中是否存在多个完全匹配的句子。当记忆库中完全匹配的记录有多个时，在“备注”列会标注“MULT”字样，提醒用户注意，如下图：

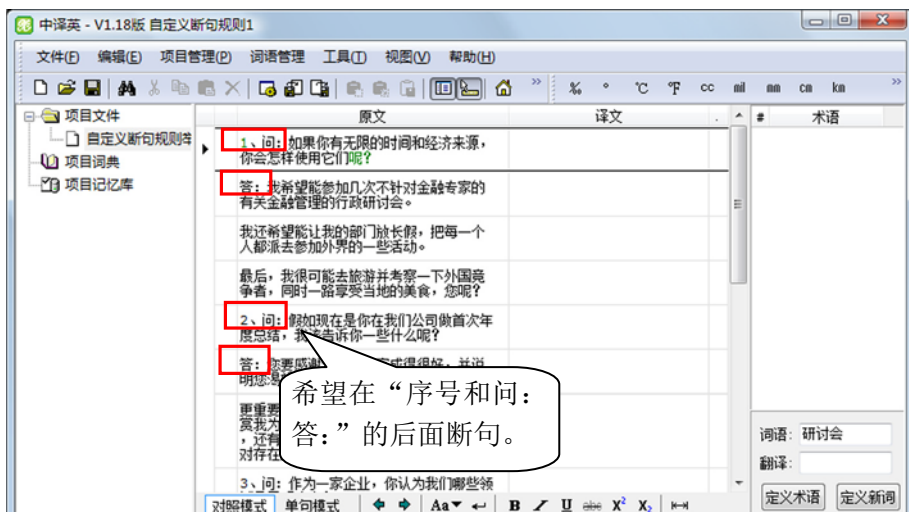




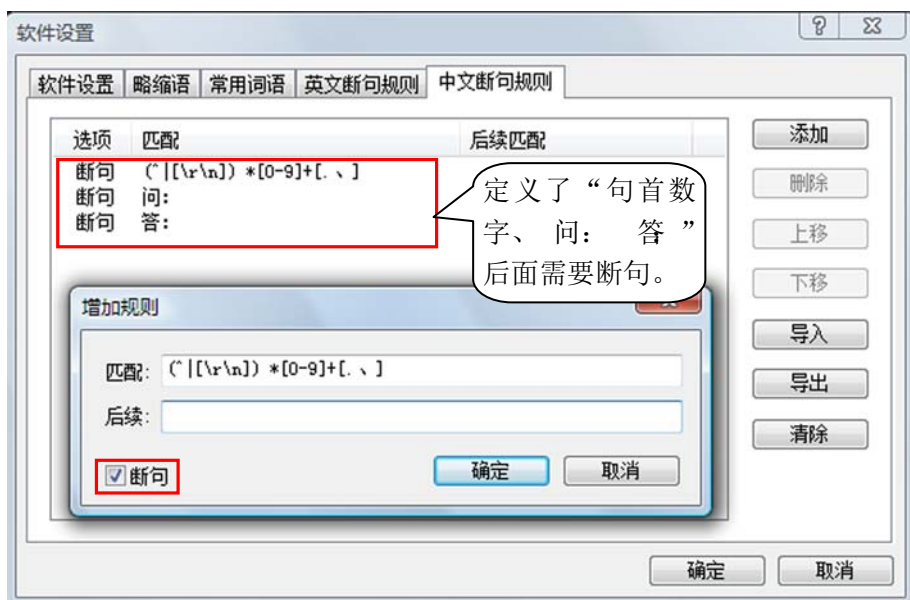
## 第七章 自定义断句规则

为了提高记忆库的匹配率，有些用户希望按“逗号”来断句，在 V1.18 版中，用户可以自行决定如何断句。

例 1：下图中：待译文章是问答方式出现的：“序号 1、2、3..... 问：..... 答：.....”的格式，若按默认的以句号断句方式，则同一行的译文也需要出现：“1、Q...”，其实我们并不希望形成的记忆库例句中含有“1、问： 答：”等等无关的词语，这些会影响该句日后的匹配率。

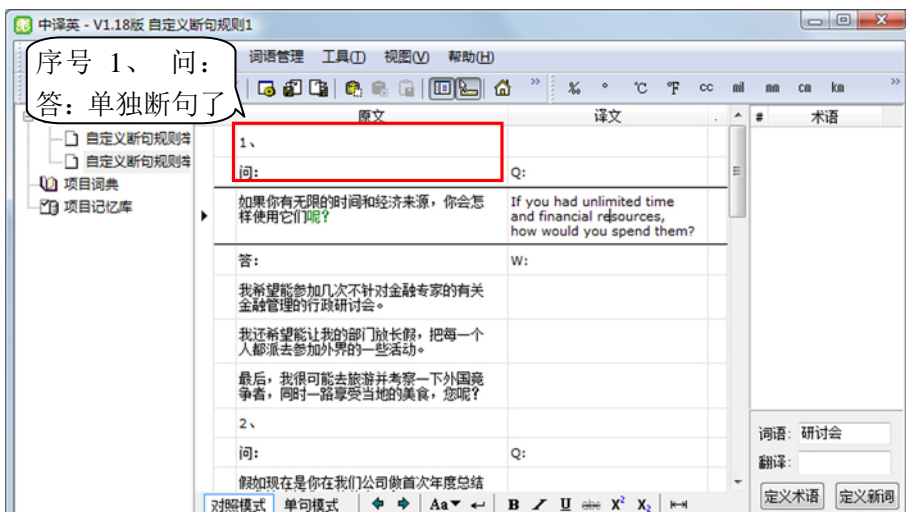


点击【工具→软件设置】菜单，由于这篇文章的原文是中文，所以点击中文短句规则页签，然后点击“**添加**”按钮，输入第一个断句规则： $(^|[\r\n]) * [0-9] + [., ]$  断句规则可以用“正则表达式”来书写，其中： $(^|[\r\n])$  表示“段首”  $[0-9]$  表示匹配 0-9 的任意数字， $[., ]$  表示数字后面是“.”或“”，例如：1、 2、 或 1. 2. ，并勾选“断句”选项，这个规则就表示：遇到句首出现的 1、2、3、或 1. 2. 3. 情况下，需要断句，点击**确定**按钮，完成第一个断句规则的定义。接着再点击“**添加**”按钮，输入第二个断句规则：“问：”后面需要断句；第三个断句规则“答：”后面也需要断句。分别见下面 3 个图：

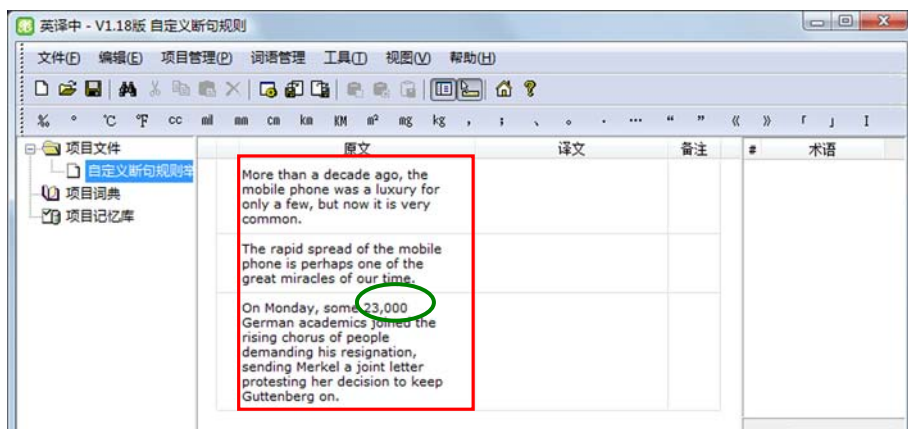


定义好这 3 个断句规则后，我们需要**重新导入**刚才的“问答文章”（**注意：软件在导入时断句，所以新定义的断句规则需要重新导入文章才能生效。**）此时软件的断句如下图：将 1、 2、 3、 等单独断成了一行， 这些数字原文译文其实是一致的，在“译文”**留空**就行了，导出译稿时雪人会自动处理的，**留空**的好处是：更快捷和该句不会被保存在记忆库中，免去了整理记忆库中的麻烦。同时“问：答：”也单独断成了一行，对这两个翻译也只需输入一次，其他的“问：答：”行，因为是“100%匹配，软件自动填写译文。”只要鼠

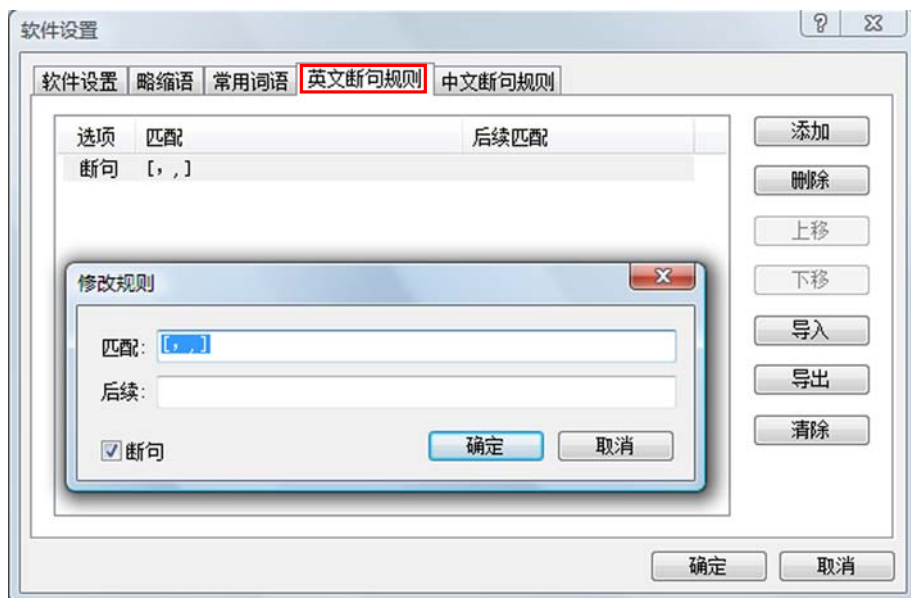
标去到该行点击一下就行了。针对每篇文章具体的特点，进行断句规则定义，显然比默认的断句更加高效！



例 2：希望按“逗号”断句，如下图中这段文字，默认是按句号断句的，若我们希望改成按逗号来断句，该如何设定断句规则呢？



由于这篇文章的原文是英文，所以点击英文短句规则页签，然后点击“**添加**”按钮，输入 [ , , ] 表示不管遇到中文形式的“逗号”，还是遇到英文形式的“逗号”，都需要断句（注意：因为在正则表达式中，“.\*() []”这些字符属于特殊字符，如果我们在断句规则中需要匹配这些符号，一定要在这些符号前加上一个反斜杠“\”，例如“\.”表示匹配符号“.”）



然后点击**确定**按钮，**重新导入**该文章，断句的结果如下图：在所有“逗号”的地方都断句了，原文 23,000 这个是数字，本不应该断句的，但现在也断句了，类似这样的情况该怎么办？



于是我们添加另外一条**不断句**的规则，如下图：在“匹配”输入框中输入[, , ]，在“后续”输入框中输入[0-9]，同时“断句”选项的前面不要勾选了，这个规则表示：当遇到中文状态下的逗号或英文状态下的逗号，若后续跟的是 0-9 的任意数字时，不需要断句，而且这个“不断句”规则必须在刚才第一个断句规则的下面，**因为当出现有多条断句规则时，最下面的规则最优先生效。**

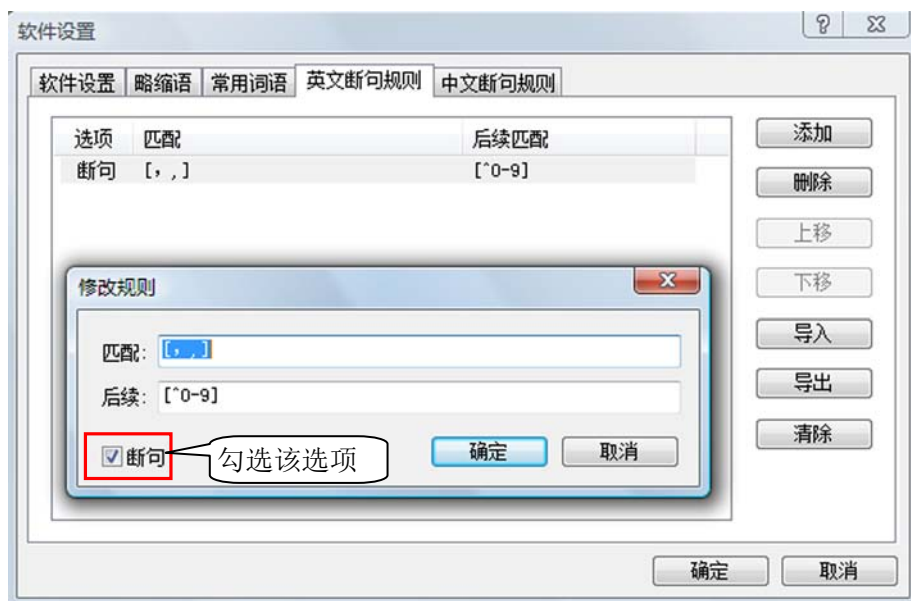


补充好不断句规则后，点击确定按钮，重新导入文章，发现在“23, 000”这个位置就不会断句了，如下图：



其实，上述的断句规则和不断句规则，可以合并在一起输入，这样使得所定义的规则更加简洁高效。如下图：在“匹配”输入框中输入[ , , ]，在“后续”输入框中输入[ ^0-9 ]（表示非 0-9），同时勾选“断句”选项，这个规则表示：当遇到中文状态下的逗号或英文状态下的逗号，且后续紧接的不是 0-9 的任意数字时就断句。如何书写高效的、随心所欲的断句规则，请大家参看有关“正则表达式”的语法规则了。

加入过多的断句规则会增加文件的导入时间，建议规则总数控制在二三十条内。



注意：自定义断句规则同样可以运用到做记忆库的“双语对齐项目”，不过，如果是双语对齐项目，中英两个方向都要作调整，中文按", "断句，英文也有按“, ”断句。否则对齐得不够好。

技巧：利用自定义断句规则，用户可以自己创建短语级的词典。即可以将一些长句拆成短语片段，然后导出 EXCEL，然后在项目词典中导入，就可以创建自己的片段短语词典了。




## 第八章 EBMT 技术实现自动替换翻译

雪人软件采用 **EBMT**(Example-Based Machine Translation 基于实例的机器翻译) 和 **TM** (Translation Memory 翻译记忆) 两种技术相结合, 充分发挥翻译记忆库的作用。

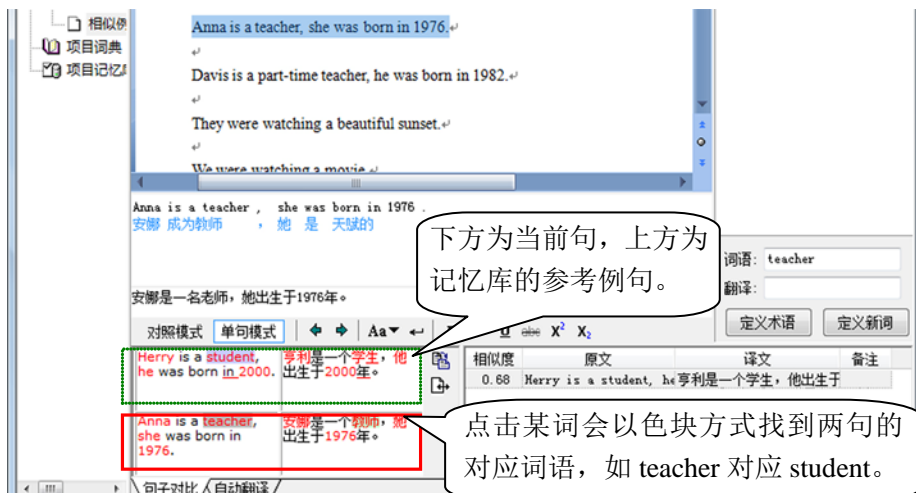
目前的 CAT 软件主要采用的是 TM 技术, 即翻译记忆 (Translation Memory, TM), TM 技术要求两句的相似度要非常高, 即基本上是完全匹配; 对两句间不同的部分, 一般只能自动替换其中的非译元素, 如数字等。所以, 记忆翻译质量的如何, 很大程度上取决记忆库的规模, 记忆库越大、覆盖面越广, 翻译的质量自然越高。但是, 记忆库不是无中生有的, 而是靠译员苦心积累下来的; 此外, 记忆库越大, 对 CAT 软件的检索、响应性能要求就越高, 对百万句对的记忆库能够做到快速响应的 CAT 软件也不多 (雪人可以轻松支持百万级的记忆库)。

而雪人的 **EBMT**(Example-Based Machine Translation 基于实例的机器翻译) 技术可以比较好的解决这一问题, 例如: 首先告诉计算机: “Herry is a student, he was born in 2000. 亨利是一个学生, 他出生于 2000 年。” 这一句的翻译。接下来的这一句 “Anna is a teacher, she was born in 1976.” 计算机能通过对上一句的类比和学习, 自动替换翻译出 “安娜是一个教师, 她出生于 1976 年。” 此外, 对两句看起来相差较多的句子, 也可通过对其中一句学习, 自动替换翻译出另外一句, 如对 “He is a student.” 和 “America is a scientic country.” 这两句, 表面看起来相似度不高, 但雪人同样可以通过对: “He is a student. 他是一个学生。” 这一句的学习, 准确翻译出 “America is a scientic country. 美国是一个科学技术发达的国家。” 这些在一般的 CAT 软件中由于匹配度不高, 会认为是两句完全不同的句子, 记忆库中的那句只能静静地沉在箱底里而不能发挥作用。下面我们来看看雪人强大实用的 “**EBMT 技术**” 和 “**清晰的句子比较**” 功能是如何充分发挥记忆库的作用!

### 一、实现一般的替换翻译

当我们翻译了第一句: “Herry is a student, he was born in 2000. 亨利是一个学生, 他出生于 2000 年。” 后, 接着翻译第二句: “Anna is a teacher, she was born in 1976.” 这时在屏幕下方的 **句子对比** 窗口中, 已经根据对第一句学习给出了 “安娜是一个教师, 她出生于 1976 年。” 的正确译文, 点击旁边的  **引用译文** 按钮, 即可将译文引用到上面。上方显示的是记忆库的参考例句, 下方显示的是当前句, 并以不同字体颜色、底

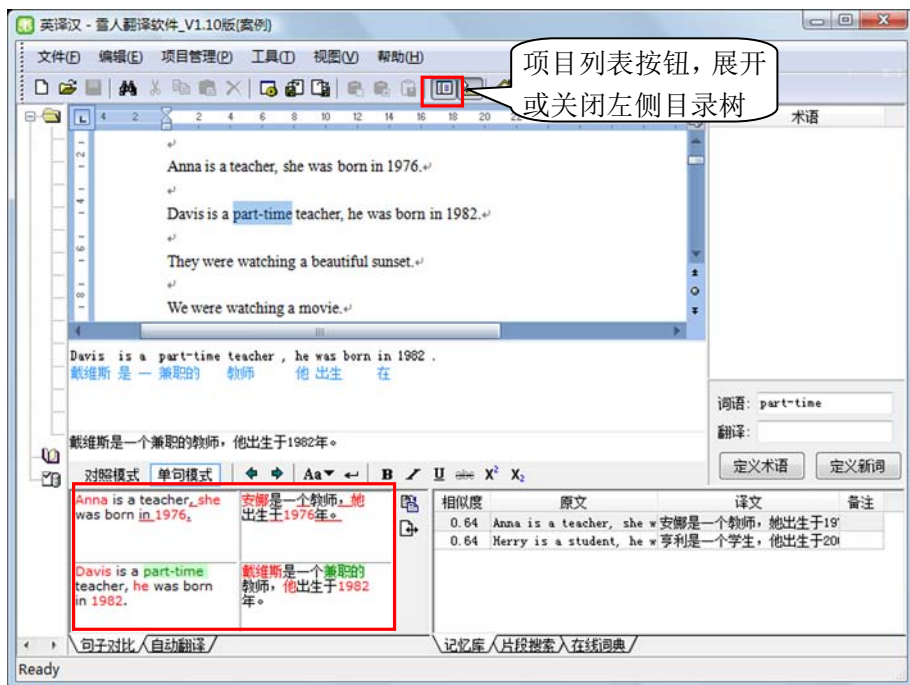
色、下划线等区分当前句子与记忆库的匹配句子的不同之处和两句之间的对应关系。在“**句子对比**”窗口中点击某词会以加亮色块方式找到与其相对应词语，如下图中，当点击“teacher”时，记忆库中的例句“student”会以色块方式加亮显示，以便译员查找两句中各词的对应关系。对于两句间不同的词语以**红色字体**显示，以便提醒译员注意。对于这两句，一般的 CAT 软件只能替换翻译数字，即将 2000 替换成 1976，不能将 Herry 替换 Anna、student 替换 teacher、he 替换 she。



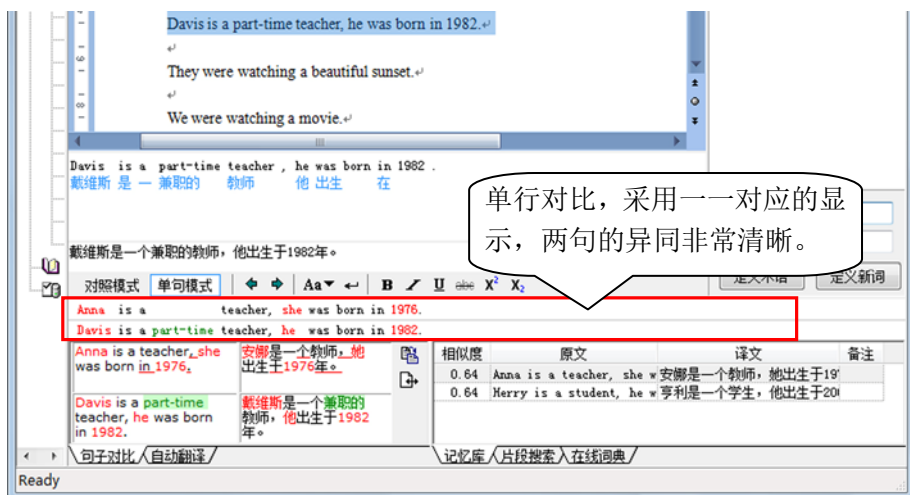
## 二、实现增词替换翻译

当翻译第三句：“Davis is a part-time teacher, he was born in 1982.” 我们看到这一句与刚才的那两句相比，结构仍旧相似，但增加了一个词“part-time”，如下图：雪人一样可根据前面的例句自动正确翻译出了译文：“戴维斯是一个兼职的教师，他出生于 1982 年。”而且对增词部分“**part-time、兼职的**”以**绿色字体**显示，以提醒译员注意该部分是否翻译正确。

另外，由于屏幕空间有限，可以点击“**项目列表**”按钮，将左侧的窗口收起来，如下图。

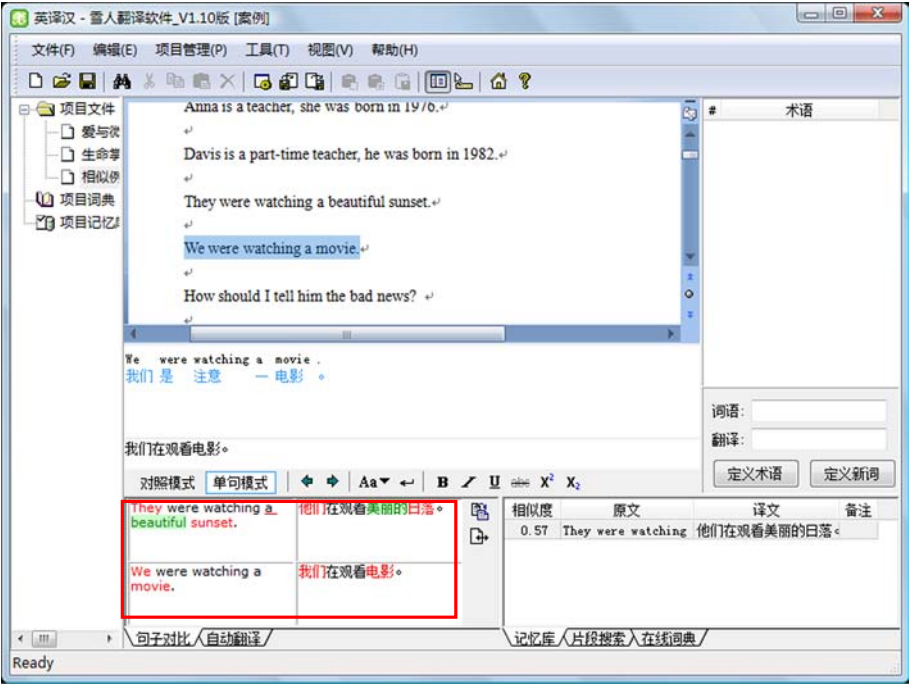


另外, 利用“单行对比”功能可以非常清晰地对比当前句子与记忆库中句子的异同。点击【视图→单行对比】菜单可以调出或关闭该功能, 界面如下: “单行对比”采用一一对应的显示, 当前句与记忆库中的参考例句的异同非常清晰, 如果句子较长, 无法显示完整, 可以在此区域中按住鼠标左键左右拖动来查看。



### 三、实现减词替换翻译

如下图：当我们翻译 “We were watching a movie.” 这一句时，软件自动从记忆库中搜索出相似度只有 0.57 的 “They were watching a beautiful sunset. 他们在观看美丽的日落。” 这一句，并以它做模板正确将 “We were watching a movie.” 翻译为 “我们在观看电影。” 而且对减词部分 “beautiful 美丽的” 以绿色字体显示，以提醒译员注意该部分是否翻译正确。这两句看似并不怎么相似，但雪人一样出色地将它译出，在一般的 CAT 软件中，记忆库中的 “They were watching a beautiful sunset. 他们在观看美丽的日落。” 是无法对当前待翻译的 “We were watching a movie.” 这一句作出贡献了。

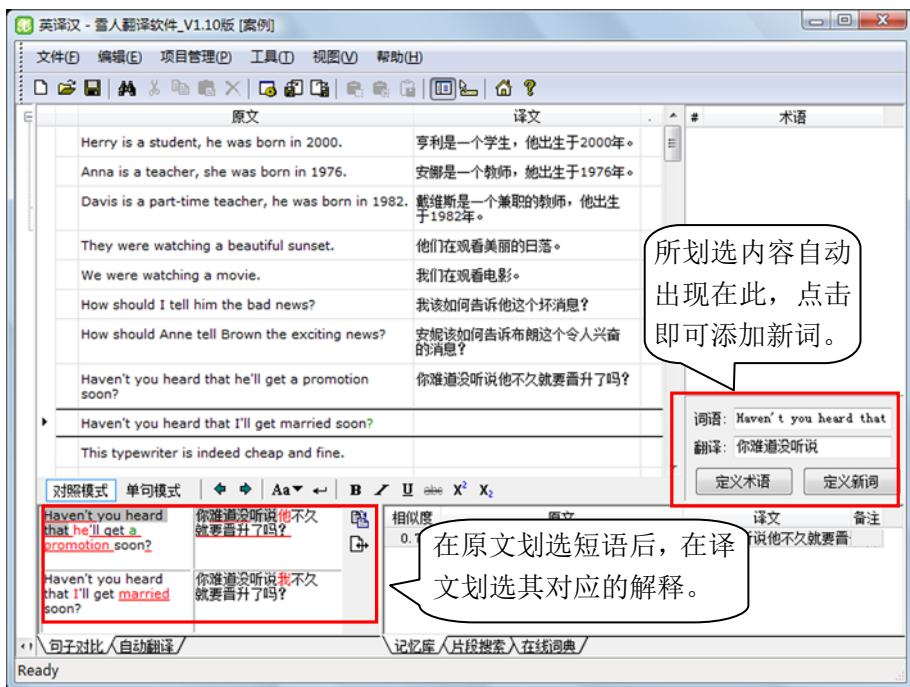



### 四、方便的短语提取

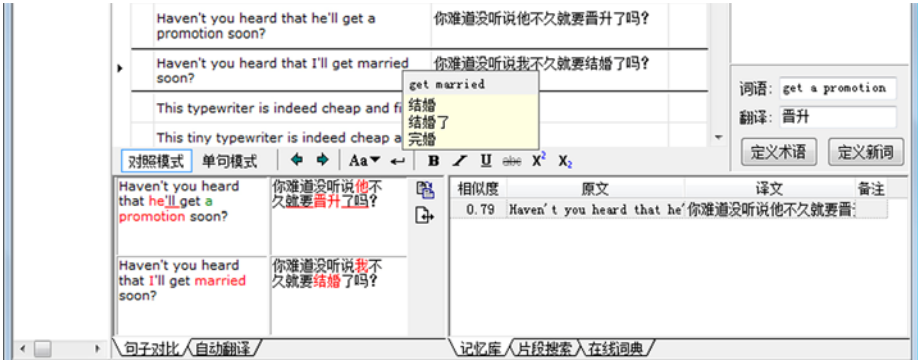
雪人可以很方便地从记忆库中提取短语、新词，而且一旦提取，立即出现在屏幕取词窗口中，以后点击即可输入；通过提取短语，还可很好地提高自动替换翻译的准确性。可谓一举多得，希望译员可以灵活运用好这个功能！如下图：当我们翻译 “Haven’t you heard that I’ll married soon?” 这一句时，从记忆库中找到了相似度 0.79 的 “Haven’t you heard that he’ll get a promotion soon? 你难道没听说他不久就要晋升了吗？” 这一句做参考，但在自动替换翻译的译文为：“你难道没听说他不久就要晋升了吗？” 这个译得并

不正确，我们需要将“晋升”自动替换为“结婚”二字，刚才那些相似度只有 0.57 的都可以正确替换，为什么这个相似度更高的却不能正确替换呢？

如下图“句子对比”窗口下方的从记忆库搜索出的那句，其原文、译文都加了红色的下划线，这是因为在我们的词典中找不到“Haven't 难道没”、“get a promotion 晋升”这样对应的解释，并以红色的下划线显示出来以示注意。这时我们先用鼠标在原文中划选“Haven't you heard that”，然后在译文中划选“你难道没听说”这几个字，刚才所划选的内容自动出现在屏幕右侧的“新词定义区”中，点击“**新词定义**”按钮或按【**Ctrl+D**】即可添加新词。同理，我们也在原文划选“get a promotion”和在译文划选“晋升”二字，并将它添加到项目词典中。



这样做了新词定义后，当鼠标再次指向该短语时，会弹出屏幕取词窗口，点击即可输入。开始软件未能正确将“晋升”自动替换为“结婚”二字是因为我们词典中缺少其相应的词汇，现在我们已经将这两个短语添加到了词典中，我们点击旁边的“重新翻译”按钮，则发现此时译文已经正确了。如下图：

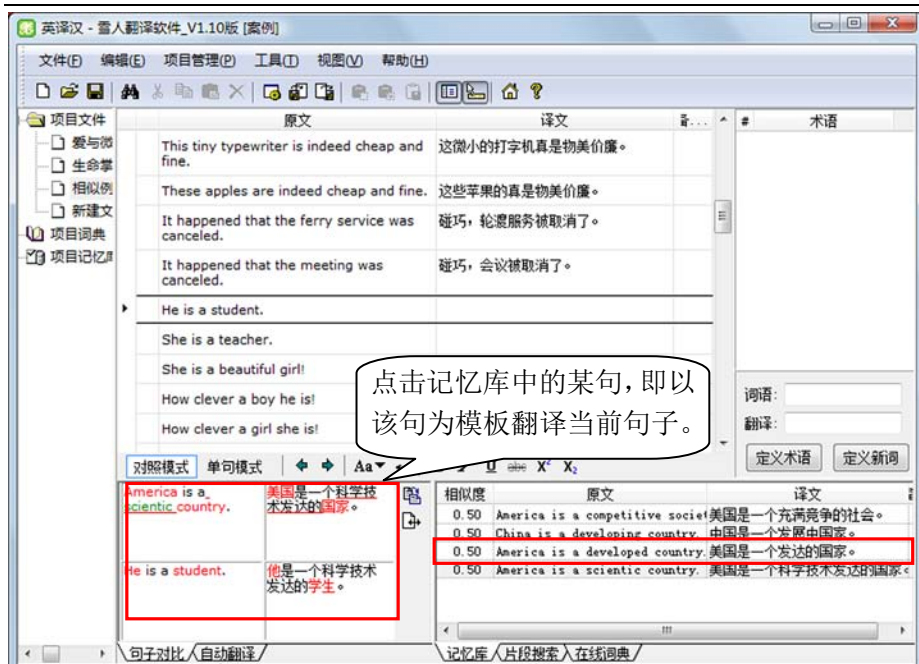



## 五、创建高质量的模板型记忆库

记忆库大固然好，但如果无法达到时，如何让有限的记忆库资源以一当十用？这不是没有可能，关键是创建高质量的“**模板型记忆库**”和“**记忆库型的用户词典**”，**让记忆库中的译文能在词典中找到对应的解释，这样的记忆库就相当强大**，对一些看似不怎么相似例句，也能根据它来翻译出正确的译文。例如：当翻译“He is a student.”这一句时，从记忆库中搜索出许多相似度只有 0.5 的例句做参考，**我们点击哪一句，软件就以这一句为模板来翻译当前句子**，如果觉得不合适，可以点击其他句子看看以哪一句为模板翻译的效果最好。一般来说，当然是匹配度越高翻译的效果越好了。这里我们选择记忆库的

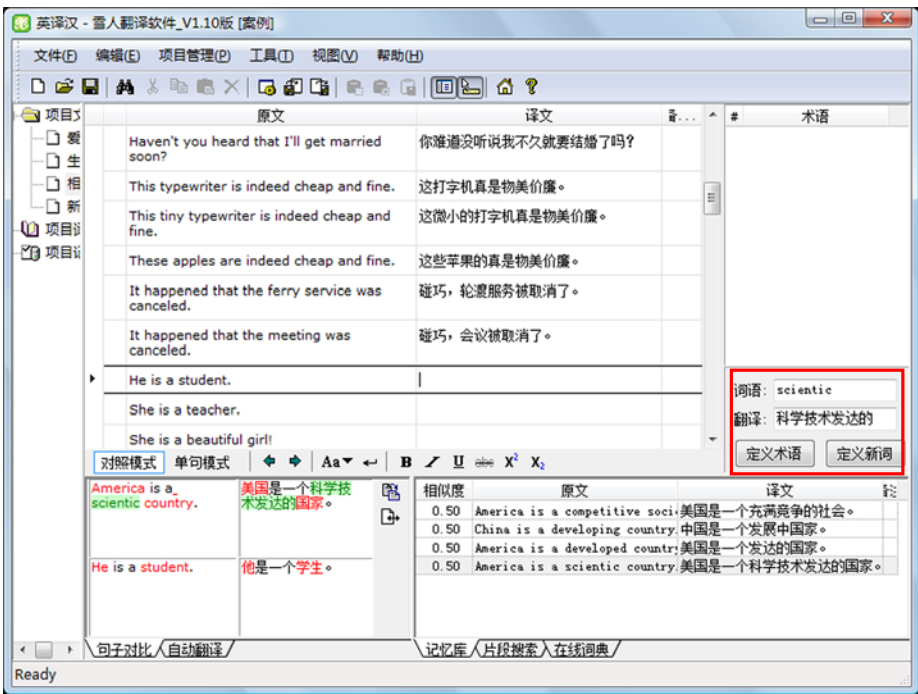
“America is a scientific country. 美国是一个科学技术发达的国家”这句为模板来翻译“He is a student.”，发现其译文是：“他是一个科学技术发达的学生”，显然不正确。软件对“America和country都能正确替换了，但对“scientific”这个词没有很好地处理，为什么呢？而且我们可以从下图中看到：“scientific”和“科学技术发达的”这两个地方都加了**红色的下划线**，这表示词典中“scientific”并没有这个意思。






这时我们划选“scientific”和“科学技术发达的”这几个字,然后将它添加到词典后,发现红色的下划线的已经消失了,表示记忆库中该句译文的所有词语在词典中都有其相应的解释。现在我们点击旁边的“**重新翻译**”按钮,则发现“He is a student.”已经正确翻译为“他是一个学生”了。如下图:一般情况下如果自动翻译译文不正确,是因为记忆库例句中有些词语的译文在词典中没有其对应的解释,而这部分又正好出现在两句中不同地方,如果词典的没有的解释出现在两句中相同的地方,还是可以正确替换翻译的。另外,如果红色的下划线出现在第一格,即当前句的原文格处,需要特别注意,此处的翻译可能不正确。

由此看来,创建一个模板型的高质量的记忆库和配套的词典,对提高自动替换翻译的准确性是相当有帮助的;对在词典中无法找到的译文解释,在记忆库的例句中会以红色的下划线显示,这时如果能对它划选进行新词定义后,那么它对以后的翻译帮助是相当大的。这个高质量的记忆库和配套词典需要译员日积月累,慢慢改进!如果做好,那么你的记忆库到时就真得可以以一当十用了。



另外，由于自动翻译取的是词典的第一个意思，如果不是需要第一个意思，可以在屏幕取词窗口中点击各词的目标意思，然后点击  重新翻译 按钮，则可以按刚才所点击的词义进行重新翻译。这也是改进准确取词的一个不错方法。

## 第九章 词语的整批替换

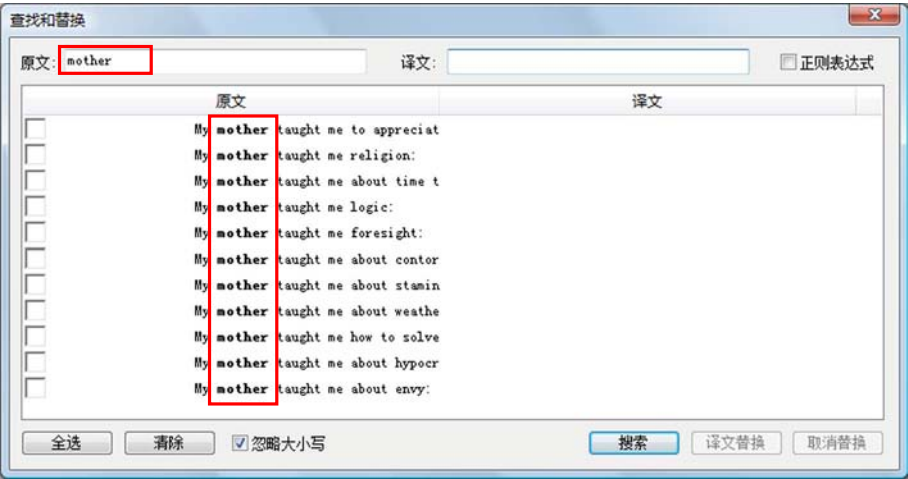
雪人的“查找和替换”功能非常强大，可以很方便地对译文、词典、记忆库进行整批的修改替换。支持“**正则表达式**”，查找方法灵活，目标结果准确，想找什么都能找到。

### 第一节 译文的整批替换

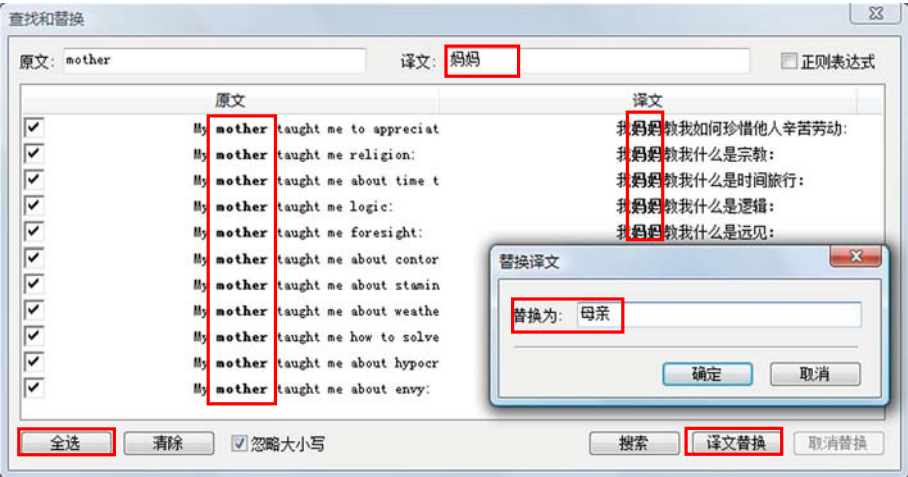
例如：我们在翻译好某篇文章后，觉得某个词之前翻译的不够准确，现在需要将它们全部找出翻译成另外一个意思，这时就需要使用“查找和替换”功能了。“查找和替换”功能可以用在原文、译文、项目词典、项目记忆库中，点击【**词语管理→查找和替换**】菜单，可以调出“查找和替换”窗口，如下图：



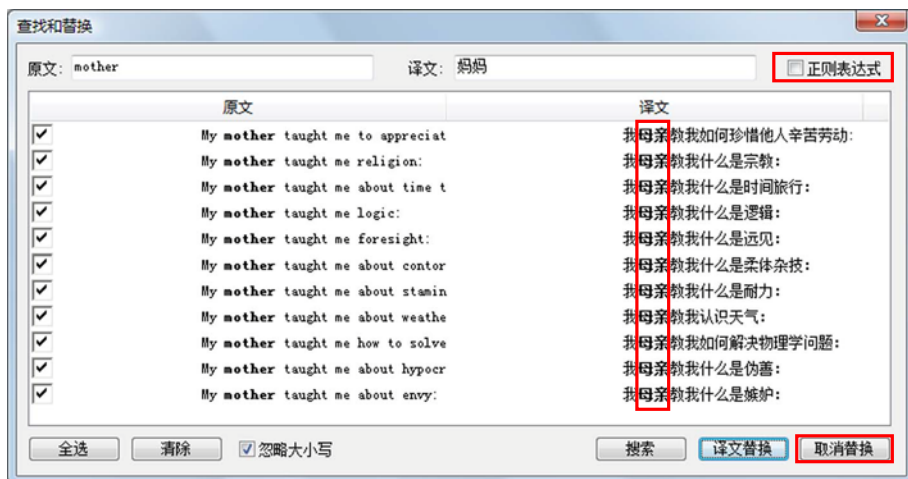
其中有“查找原文输入框”和“查找译文输入框”，还有“是否使用正则表达式”和“是否忽略大小写这两个选项”。如“mother”之前我们翻译成“妈妈”，但后来觉得译成“母亲”会更好些。此时可在“查找原文输入框”输入 mother 这个单词按回车键后，雪人即刻在原文中查找所有含“mother”这个单词的句子，并将“**mother**”居中加粗显示出来。



然后我们在对应的译文中输入其译文“妈妈”，按回车后，雪人即刻显示各句相应的译文，同时“妈妈”这两个字也被居中加粗显示，如下图：点击下方的“**全选**”按钮（或点击前面的选择框，逐个选择哪些需要替换），然后点击屏幕下方的“**译文替换**”按钮，并输入替换的译文，如：“母亲”。



点击“**确定**”按钮后，则所有“mother”的译文就会被替换成了“母亲”。如下图：此时，若发现替换的不对，按“**取消替换**”按钮，还可以回复到替换前的状态。若按“**搜索**”按钮继续查找其他词或退出该窗口，则替换就不可逆了。



## 第二节 正则表达式

上述只是对简单的词汇进行查找、替换，雪人还有个更加强大的查找功能，就是对“**正则表达式**”的支持，利用它可以实现复杂条件下的精确查找，使得查找随心所欲！

### 一、正则表达式理论知识

正则表达式就是用于描述查找符合某些复杂规则的字符串的代码。类似于 Windows/Dos 下用于文件查找的通配符，如\*.doc 可以查找某个目录下的所有的 Word 文档，只不过正则表达式更能精确实现你的查找目标，同时也会更加复杂些，这里只简单介绍几个常用的元字符，若需更加深入地学习正则表达式的使用，请参考“正则表达式”的相关书籍。

\b——是正则表达式的一个特殊代码，代表单词的开头或结尾。有了它，当你搜索“is”这个单词时，this、his、miss 等就不会出现，写法：\b 单词\b，如：\bis\b，表示仅仅搜索 is 这个单词，其他含有 is 这两个字母的单词不要出现。

.——是一个很有用的元字符，它可以匹配除了换行符以外的任意字符。

\*——同样也是一个很常用的元字符，不过它代表的不是字符，也不是位置，而是数量；它指定\*前边的内容可以连续重复使用任意次以使整个表达式得到匹配。

. \*——.\*连在一起就意味着任意数量的不包含换行的字符。例如：\bis\b.\*\bhappy\b 的

意思是: 搜一个单词 is, 然后是任意个任意字符(但不能是换行), 最后是 happy 这个单词。  
`^`——匹配字符串的开始。( `^` 是和数字 6 在同一个键位上的符号)

`$`——匹配字符串的结束。

`\d`——匹配数字。

`\s`——匹配任意的空白符。

`\w`——匹配字母或数字或下划线或汉字

## 反义:

有时需要查找不属于某个能简单定义的字符类的字符。比如想查找除了数字以外, 其它任意字符都行的情况, 这时需要用到反义:

`\W` ——匹配任意不是字母, 数字, 下划线, 汉字的字符

`\S` ——匹配任意不是空白符的字符

`\D` ——匹配任意非数字的字符

`\B` ——匹配不是单词开头或结束的位置

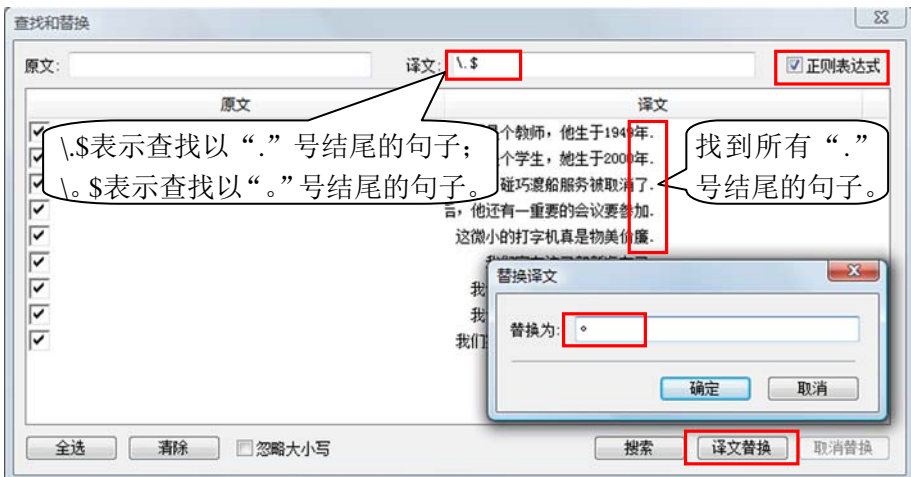
`[^x]`——匹配除了 x 以外的任意字符

`[^ness]`——匹配除了 ness 这几个字母以外的任意字符

## 二、正在表达式实际应用

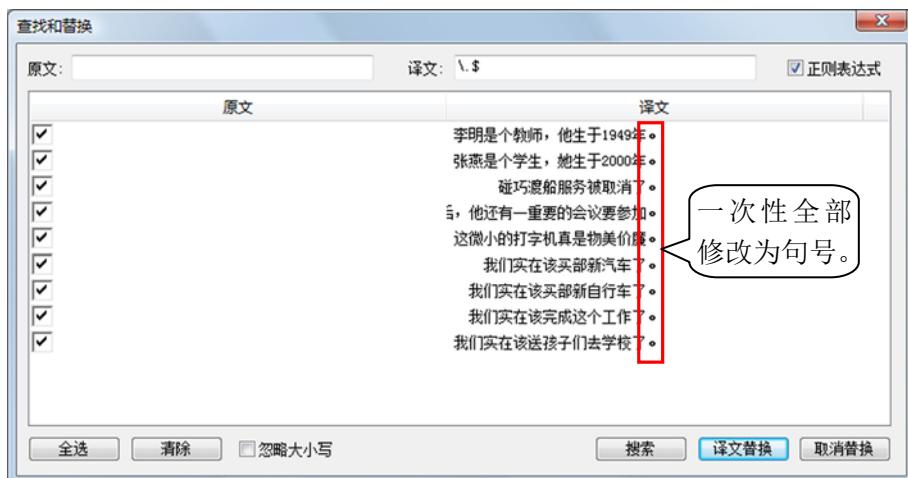
### 1. 将句尾的“.”号全部替换成“。”号

在英译中时, 经常出现用英文的“.”号来做句子的句号。如何将它们全部找出一并替换成“。”号呢? 下图中, 我们先勾选右上角的“正则表达式”, 勾选则表示启用该功能。然后在译文输入框中输入: `\.$` 表示查找以“.”号结尾的句子(若 `\.$` 则表示查找以“。”号结尾的句子), 按回车后即可搜索出所有以“.”号结尾的句子。



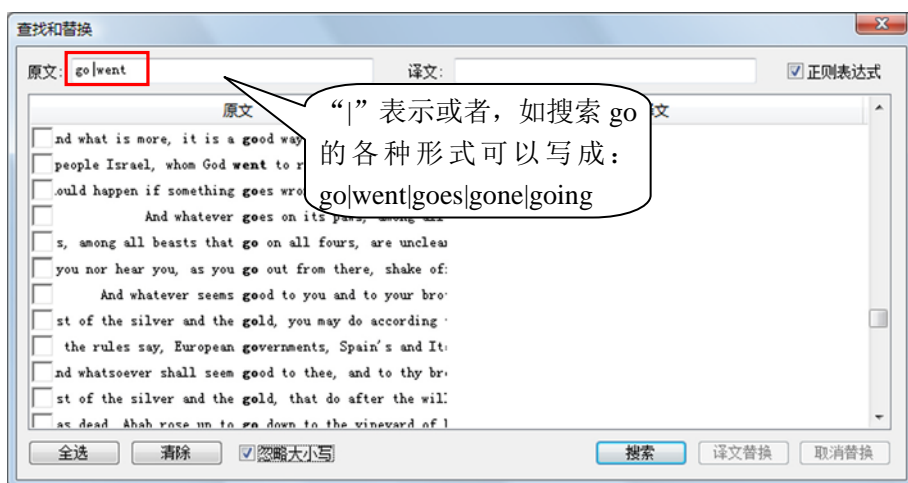


全选这些句子后，点击“**译文替换**”按钮，并输入“。”，就可以将它们一次性全部修改为“。”了。替换结果如下图：



## 2. 搜索多个词

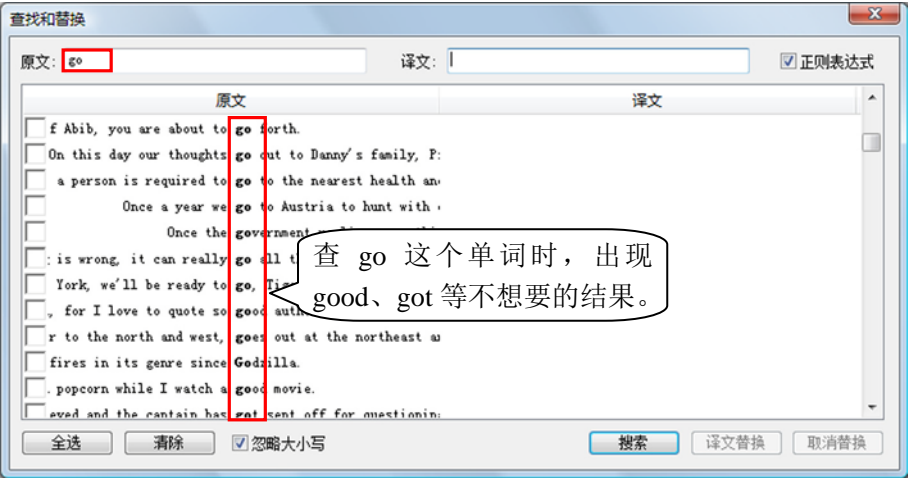
输入 go|went 表示搜索含 go 或 went 的单词，例如：go|went|goes|gone|going，可以搜索出 go 的各种变形词。



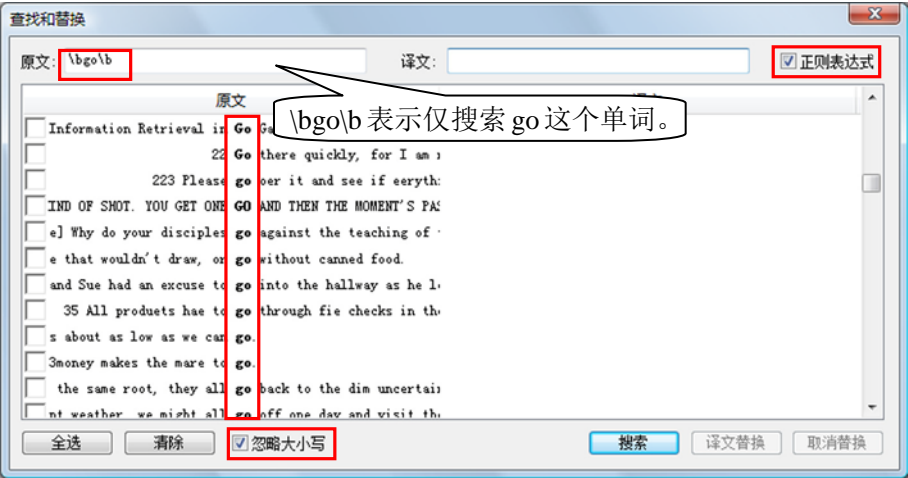
## 3. 精确查找单词

例如想查找“go”这个单词，在原文输入框中输入“go”时，查找到的结果除了有“go”这个单词外，还出现了许多其他含 go 这两个字母的单词，例如 good、got 等等，这些当然

不是我们想要的结果。如下图：

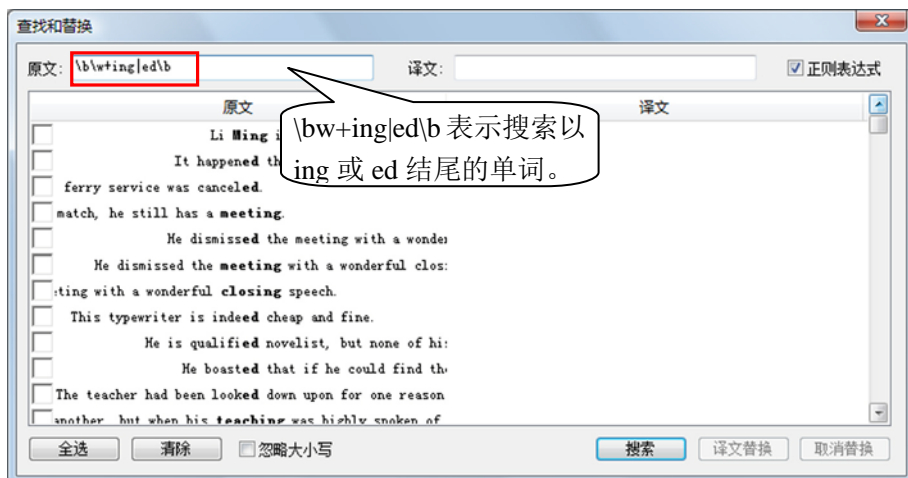
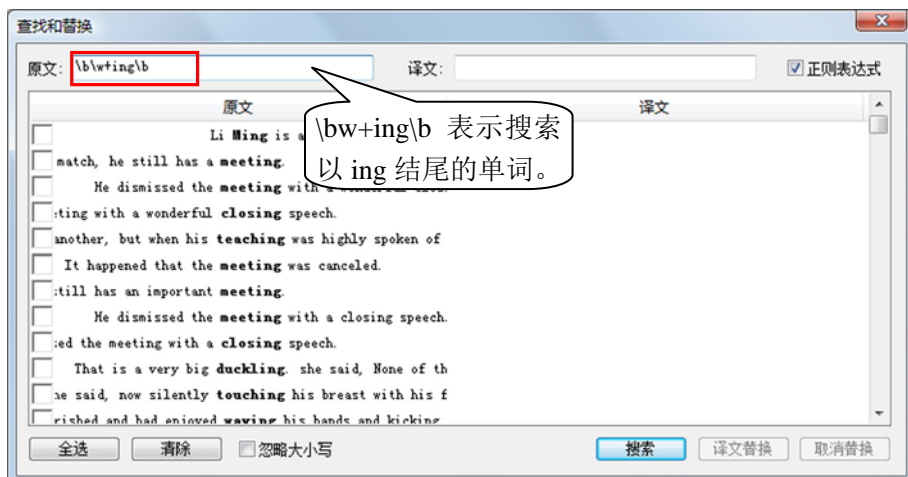


但若是输入的查找条件写成：“\bgo\b”表示只查找 go 这个单词，同时勾选了屏幕下方的“忽略大小写”选项，则搜索的结果就只是出现 go 这个单词了，还包括 Go、G0 大小写的变化形式，而 got、good 等无关的词汇则不再出现。如下图：



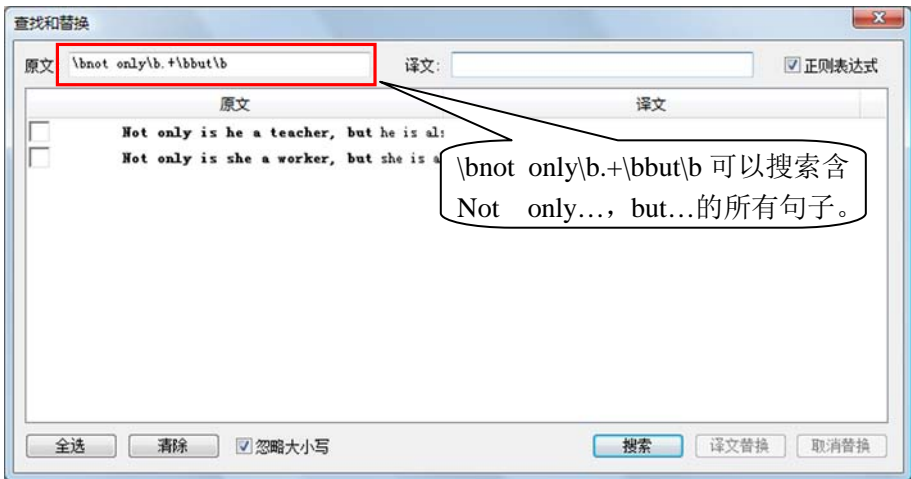
\b\w+ing\b            表示搜索以 ing 结尾的词。搜索结果如下图：

\bw+ing|ed\b        表示搜索以 ing 或 ed 结尾的单词。搜索结果如下图：



#### 4. 对间隔短语的搜索

利用“正在表达式”的强大搜索功能，可以搜索出中间带省略号的复杂短语，例如希望搜出所有含 Not only..., but...的句子，只需要在下图的搜索框中输入：\\bnot only\\b.+\\bbut\\b 即可实现这一需求，搜索结果如下图：



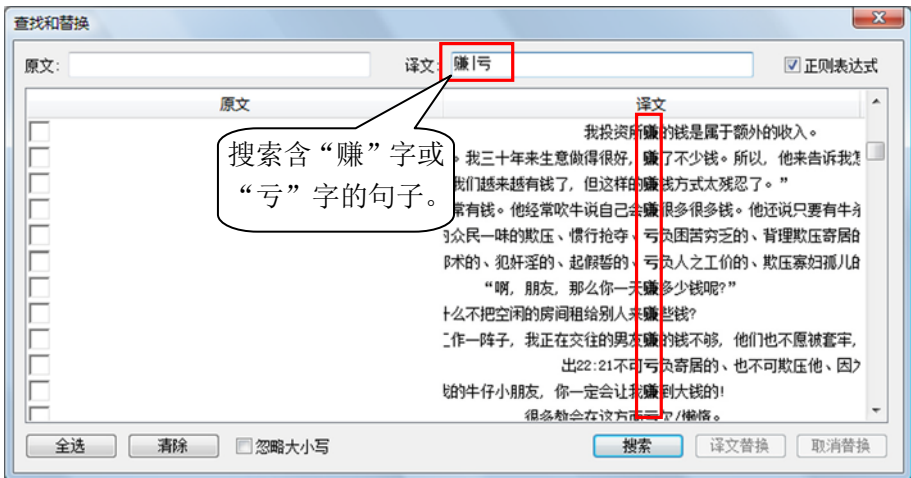
\\bnot only\\b.+\\bbut \\b 搜索固定空格数的 Not only..., but...短语。

若 not 和 only 之间是不固定空格数，如 not ... only..., but...also...则应该写成：

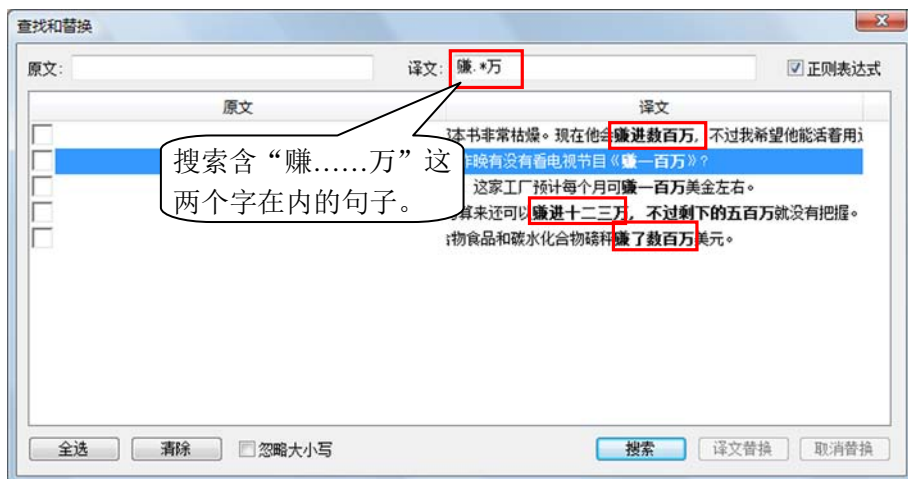
\\bnot +only\\b.+\\bbut +also\\b

## 5. 中文搜索

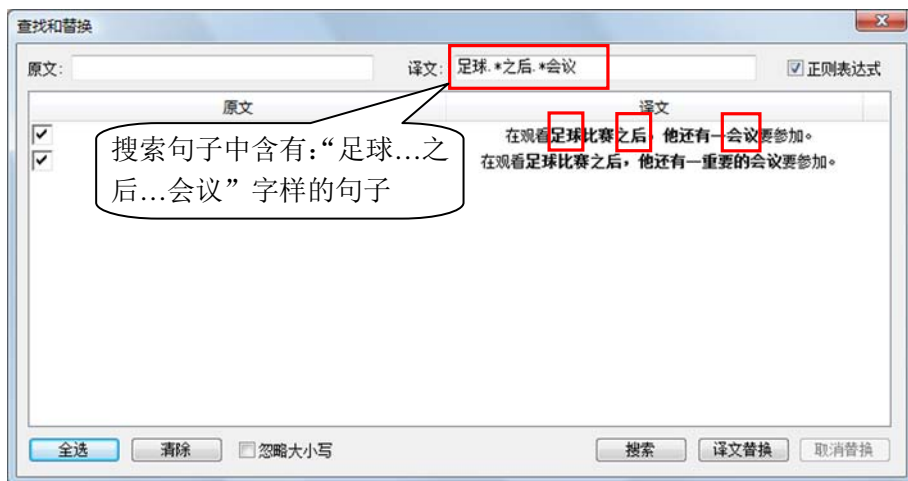
“正则表达式”对中文的搜索也是非常方便的，例如输入：赚|亏 可以搜索包含“赚”字或“亏”字的句子。搜索结果如下图：



若输入：赚.\*万 可以搜索含“赚……万”的句子，搜索结果如下图：有赚进数百万、赚一百万、赚进十一二万等等。



若输入: 足球.\*之后.\*会议 表示搜索句子中含有: “足球……. 之后……会议”字样的句子。如下图



正则表达式的使用有标准的语法, 更多关于正则表达式的高级运用, 请用户自行查阅相关的正则表达式的书籍。

### 三、自动记录查找条件

为了使一些常用的查找条件不用总是重复输入, 软件增加了自动记录“查找条件”功能, 点击【词语管理→查找与替换】菜单, 在“查找与替换”窗口的右下角增加了一个“更多”按钮, 点击该按钮, 可展开窗口, 其中有“历史记录”和“常用记录”两个页签, “历史记录”页签用于自动记录用户曾经输入过的查找条件, 它仅对当前的翻译项目有用, 最

查找和替换

原文:  译文:  ☒ 正则表达式

☐ ☐ ☐ ☐ ☐ ☐ ☐ ☐ ☐ ☐

原文 译文

(真棒)等词语来表达对你的感谢。  
 现在更喜欢用cheers来表达感谢。  
 住、更加口语化的词来表示感谢。  
 手来表示感谢,而不说thank you。  
 用一个友好的动作来表示效果好。  
 英国网上礼品店to You开展的。  
 们还是会用其他方式来表达感谢。  
 你的方式是友好礼貌的就可以了。  
 英国人应该多学学如何表达感谢。

全选 清除 ☐ 忽略大小写 搜索 译文替换 取消替换 更多...

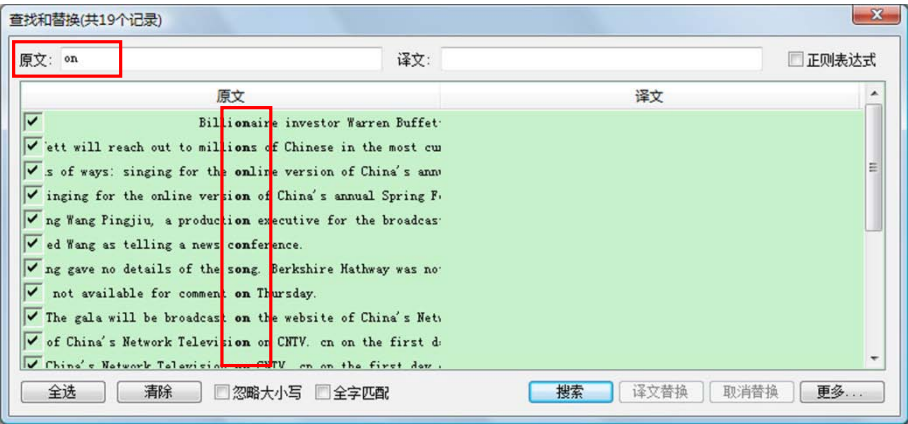
序号	原文	译文	选项	备注
1		\. \$	正侧表达式	搜索以.号结尾的句子
2		\. \$	正侧表达式	搜索以.号结尾的句子
3	\blw+ing\b		正侧表达式	搜索以ing结尾的单词
4	\bwing+ed\b		正侧表达式	搜索以eng结尾的单词
5	\bnot only\b.+but\b		正侧表达式/大小写无关	搜索含not only...but...的句子
6				

历史记录 常用记录



## 四、全字匹配

在“查找和替换”窗口中，例如搜索“on”，则那些包含“on”的单词也会显示出来，如 song、online 等等也会出现在搜索的结果中，如下图：

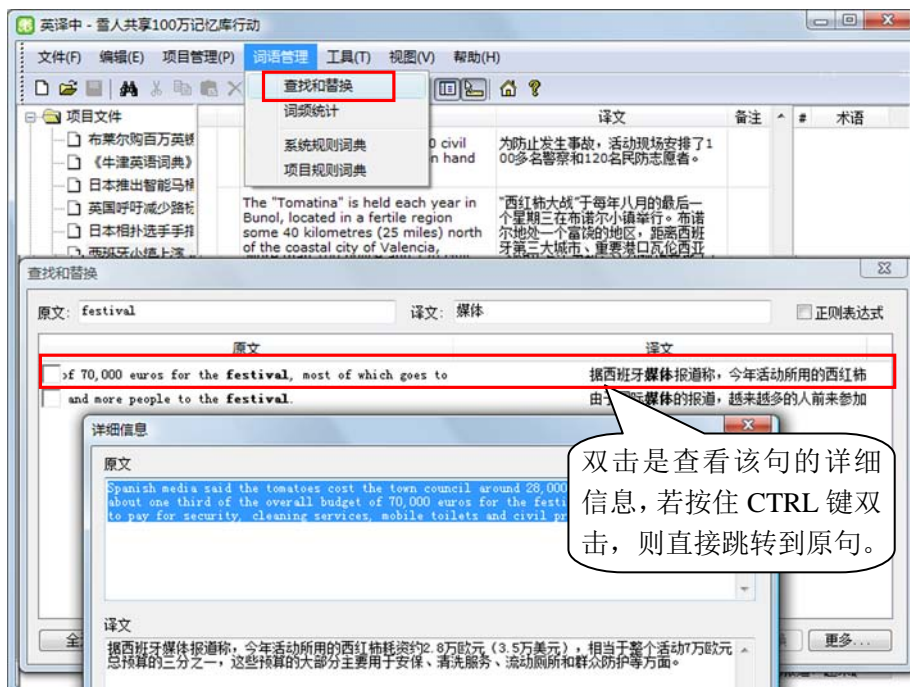


如果勾选窗口下方的“全字匹配”选项后重新搜索，仅仅出现是以“on”为单词的句子，其他单词中包含“on”的句子则不会再被列出来了，搜索的结果会更加清晰、准确。如下图：（注：使用正则表达式时自动忽略该选项）



## 第三节 对查找的结果直接定位到原处

点击【词语管理→查找和替换】菜单，在查找和替换窗口中，若查找到的结果比较长，不能完整查看时，双击该句，会弹出“详细信息”窗口，用于完整查看，如下图。



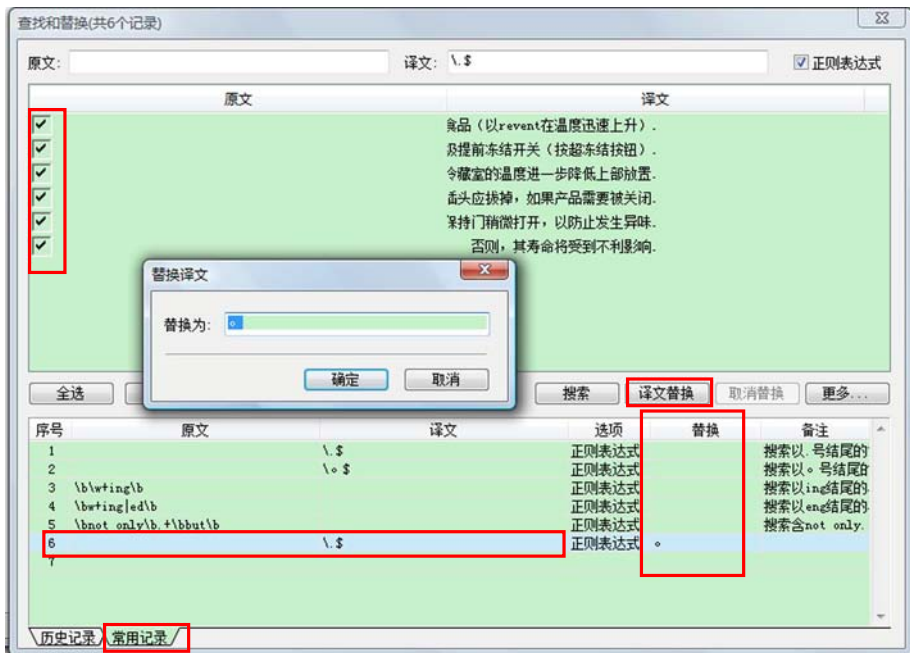
若按下 **CTRL** 键双击，则可以直接跳转到原句。如下图。



#### 第四节 设置默认的替换内容


在 V1.25 版的查找替换中，将搜索结果的内容默认为全部勾选状态，方便用户对查找结果进行其他操作。此外，“常用记录”还支持了设置默认的替换内容，用户可以将一些经常用到的查找、替换操作添加到常用记录中，并在“替换”列中设置默认的替换内容，以后只要点击“替换译文”按钮就可以瞬间完成，省去了再次输入查找条件、输入替换译文等的操作。例如：经常需要查找译文中不正确的标点符号，将“句号”替换成“逗号”，我们将“查找条件”及“替换的内容”在“常用记录”页签中输入，如下图：以后只要点击“替换译文”按钮就实现将所有以“。”结尾的句子修改成按“，”结尾了。用户可以根据实际需要，灵活设置查找条件及替换内容，例如用在繁体和简体转换时，一些固定词语的替换上，是非常实用、高效的。

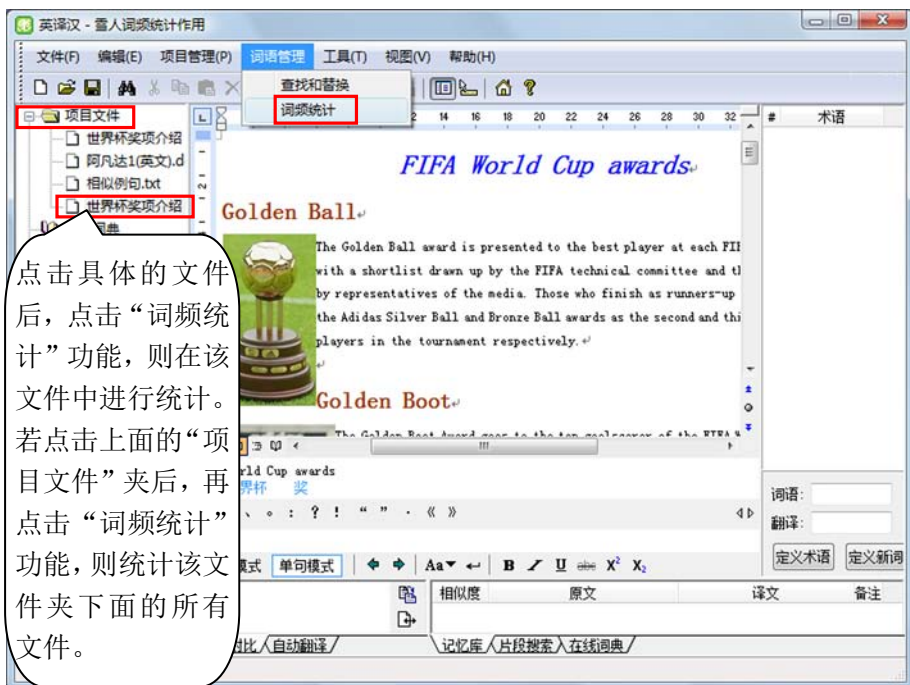
注意：如果需要将常用记录页签中的内容导出到 EXCEL，则应该避免以“=”字符开头，否则导出会报出错。



## 第十章 词频统计（词语提取）

### 一、从待译文章中提取词语

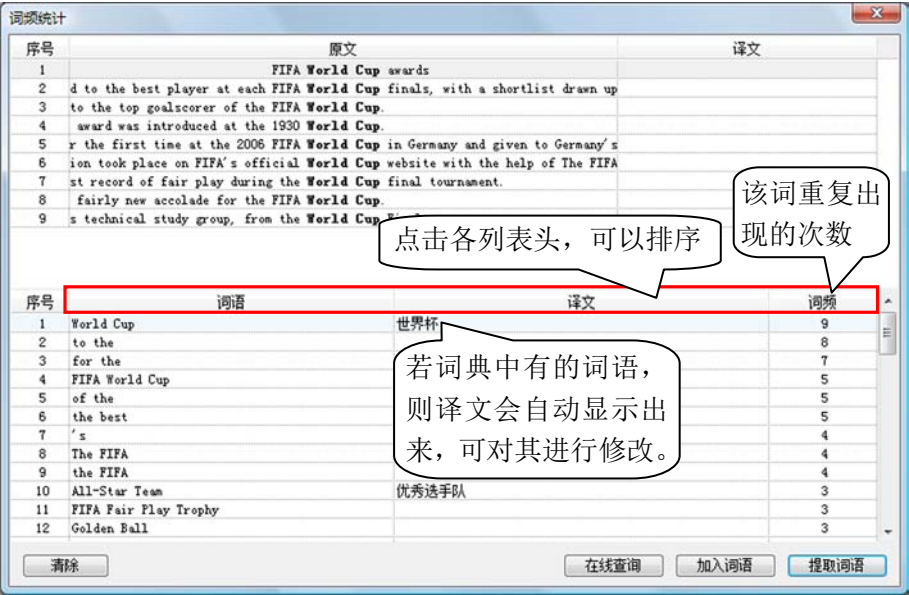
利用雪人的“词频统计”功能可以统计出文章中的高频词语，若在翻译之前对它们进行准确的定义，将会对日后的翻译工作起到事半功倍的作用。雪人支持对单个文件进行词频统计，亦支持同时对多个文件进行词频统计，点击具体的某个文件后（如：“世界杯奖项介绍”这个文件）后，点击【词语管理→词频统计】菜单，则仅统计“世界杯奖项介绍”这篇文章中的词频情况。若点击上面的“ 项目文件”后，再点击“词频统计”功能，则统计该文件夹下面的所有文件的词语重复情况。



例如：我们想在翻译之前对“世界杯奖项介绍”这篇文章的高频词语进行预先统计、定义，则点击屏幕左侧的“世界杯奖项介绍”，然后点击【词语管理→词频统计】，弹出的词频统计窗口如下图：

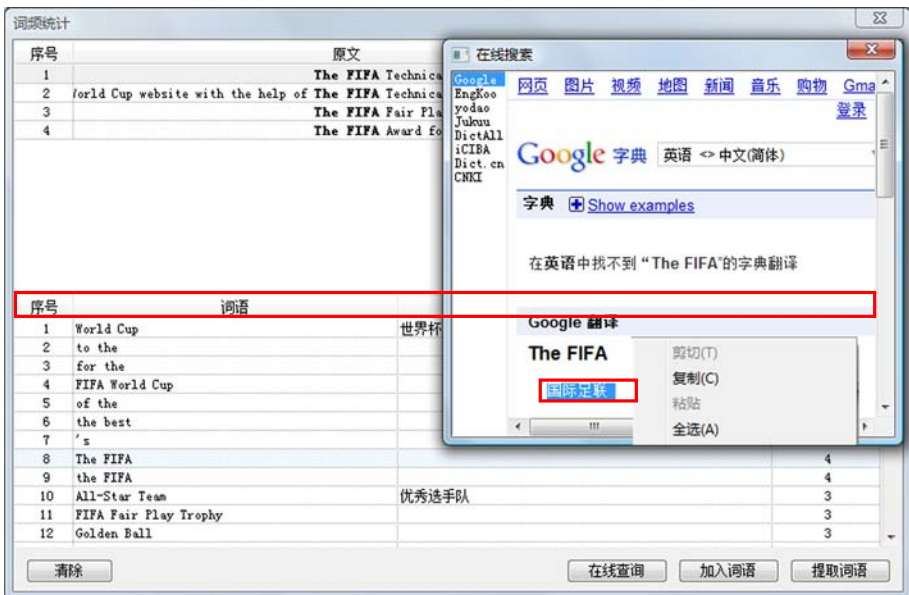


点击右下角的“**词语提取**”按钮，在弹出的对话框中输入词语出现的最少频率和长度，具体数值由用户自行决定，例如在“最少出现词频”中输入 5，则表示必须在文章中重复出现 5 次以上的词语才会被提取出来。按“**确定**”按钮后，提取结果如下图：

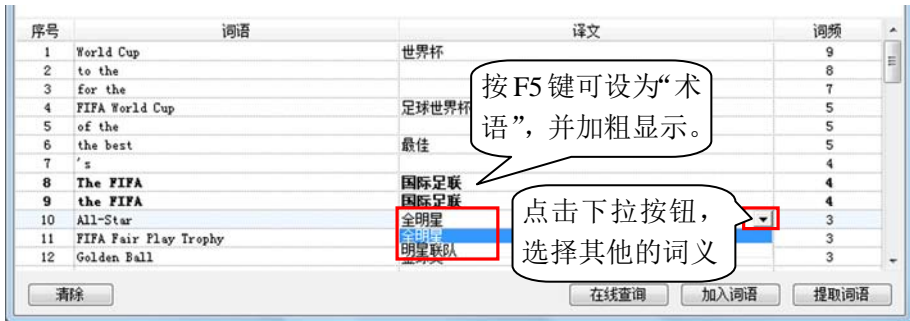


中部“词语、译文、词频”各列的表头，点击它可以排序，若词典中有的词语，则译文会自动显示出来，若认为词义不准确，用户可在“译文”列中对其进行修改。上述例子是一篇简短的文章，若是一些大型的技术性文章，词频数值可能到达几百次、甚至千次以上，对这些高频词进行准确的定义，会使后续翻译工作更加高效！

若对某高频词不知该如何翻译，雪人在此提供了非常方便的“在线查询”功能，点击“**在线查询**”按钮，即刻从在线词典中查到结果。例如下图中，若对“The FIFA”不清楚该如何翻译？则点击“**在线查询**”按钮，即刻从 Google 中找到了“The FIFA 国际足联”，然后利用【右键菜单→复制】和【粘贴】功能，将词义复制下来并粘贴到译文处，如下图：

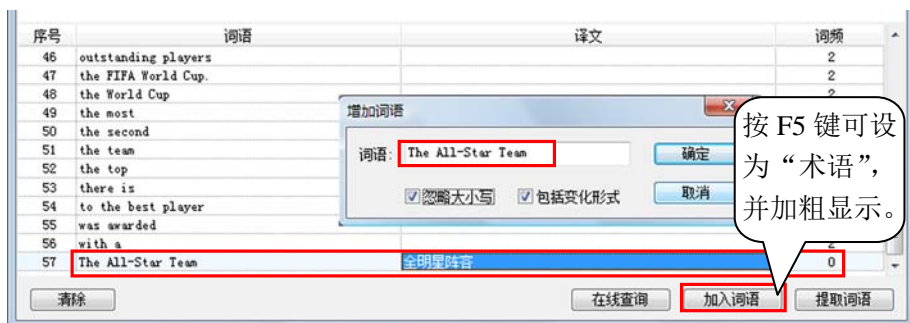


按【F5】键，可将该词语设为术语，并**加粗**显示。若词典中存在的词语，由于存在一词多义的情况，点击其相应的译文，在旁边会出现一个▼下拉按钮，可以更换其他的词义。

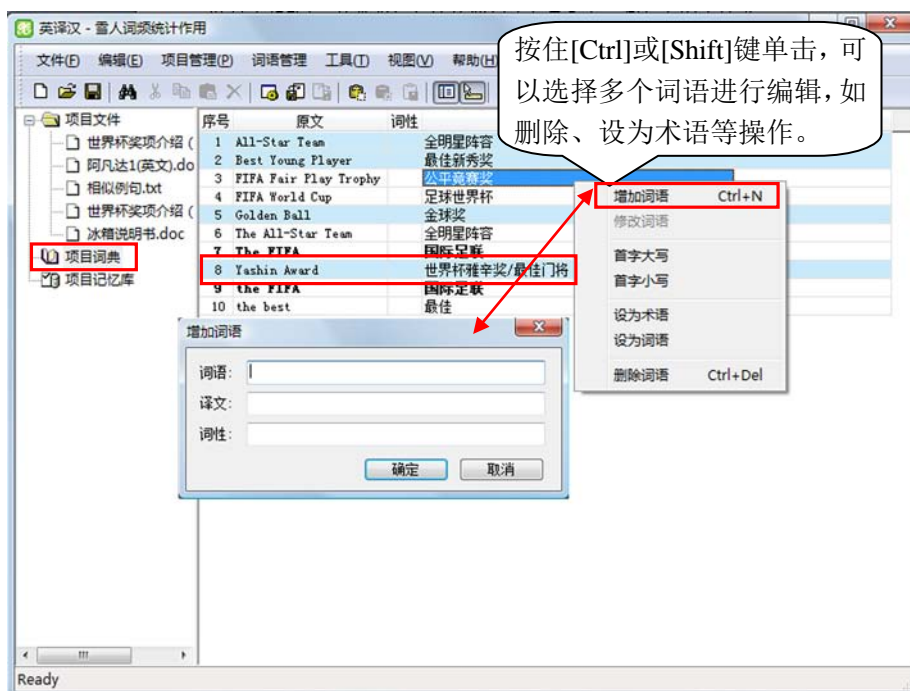




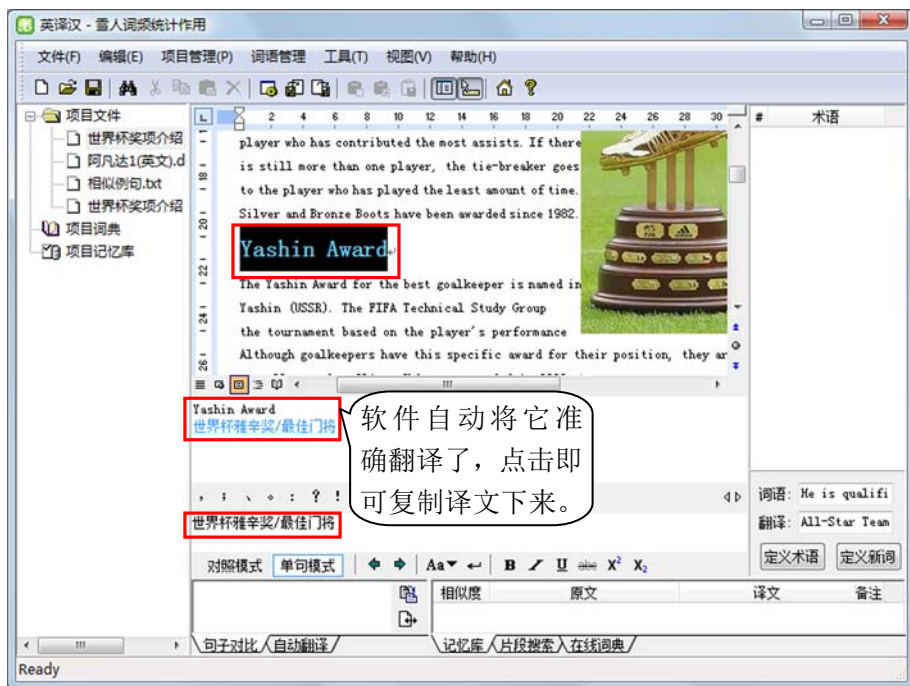
有时软件自动提取词语的断句、搭配不一定符合要求，这时用户可以点击“**加入词语**”按钮，自己增加一些新词（可参考已经提取的词语，利用【Ctrl+C】和【Ctrl+V】复制粘贴过来修改），软件将增加的这些词放在最后面，例如下图中的“The All-Star Team 全明星阵容”是自己添加的。






定义好高频词的词义后，退出“词频统计”窗口，然后点击屏幕左侧的“**项目词典**”，可以看到刚才我们定义的这些高频词，这些词将会作为词典的默认第一词义出现在译文中，在**项目词典**窗口中，按住【Ctrl】或【Shift】键单击，可以选择多个词语进行编辑，如删除、设为术语或整批设为“首字母大写、小写”等操作，还可以直接增加新词，具体功能见下图的右键菜单中所列：



例如上图中，在词频统计中发现“Yashin Award”在文章中多次重复出现，对它进行了“世界杯雅辛奖/最佳门将”的定义后，当翻译时遇到该词语时，软件已经根据所定义的词义对它准确翻译了，点击即可将译文复制下来，对高频词的预先定义，将给后面的翻译工作带来很大的方便。



## 二、从记忆库中提取词语

雪人的“词频统计”功能不但可以在翻译的文章中提取高频词，将它们定义到  **项目词典** 中，还可以对记忆库中的句子进行词语提取，统计出记忆库中的高频词语，将它们也保存到  **项目词典** 中。方法是：首先点击屏幕左侧的  **项目记忆库**，然后再点击【**词语管理**→**词频统计**】菜单，则就可以在记忆库中进行词语提取了。例如下图中：从记忆库中自动提取到“in accordance with”这个短语的出现频率是 601 次。

词频统计

序号	原文	译文
1	ations to the creditor	in accordance with the provisions of this
2	a new collegial panel	in accordance with the procedure of first
3	heir litigation rights	in accordance with the law, observe the p
4	people's court shall,	in accordance with the procedure prescrib
5	es is always regulated	in accordance with the imagination. Man
6	), while still being	in accordance with rule 3, and thereby si
7	this CONTRACT strictly	in accordance with the provision hereof.
8	reduction or drawback	in accordance with the relevant provision
9	imposed on shipowners	in accordance with this Convention, regar
10	ficial weekly rest day	in accordance with the customs of Germany
11	mum (40) hours per week	in accordance with the relevant laws and
12	mpany shall be handled	in accordance with the foreign investment
13	6.1 and addenda issued	in accordance with Clause 8 shall be deem
14	ich shall be conducted	in accordance with the commission's arbit
15	losing date of tenders	in accordance with Clause 8 and any minut
16	tender meetings issued	in accordance with Clause 16 of these Ins
17	Leased Unit by Party B	in accordance with this Lease to Party A
18	of the Product and not	in accordance with the provision of the m
19	ct the suitable grease	in accordance with the lubrication chart.
20	if it has been climbed	in accordance with the existing rules and

从记忆库中自动提取词语

清除

在线查询

加入词语

提取词语

序号	词语	译文	词频
7	Abstract: The		602
8	in accordance with	依据	601
9	...		593
10	as well as		586

三、查看长句

当遇到比较长的句子时，无法看到完整的句子，此时**双击**某长句，可弹出“详细信息”窗口，这样便可以看清长句的所有内容了。

词频统计

序号	原文	译文
1	FIFA World Cup awards	世界杯奖项介绍
2	d to the best player at each FIFA World Cup finals, with a shortlist drawn up	世界杯金球奖 世界杯金球奖用来颁发给每届世界
3	to the top goalscorer of the FIFA World Cup. The award was introduced at the	世界杯最佳射手将获得金靴奖，于1930年
4	award was introduced at the 1930 World Cup.	世界杯最佳射手将获得金靴奖，于1930年
5	the first time at the 2006 FIFA World Cup in Germany and given to Germany's	2006年德国世界杯上首次颁发了佳新秀，

双击可弹出详细信息窗口，方便查看长句的原文、译文。该窗口仅供查看，不适用于修改，即使修改亦不会反映到原文处。

详细信息

原文

The Best Young Player award was awarded for the first time at the 2006 FIFA W  
Germany and given to Germany's Lukas Podolski.

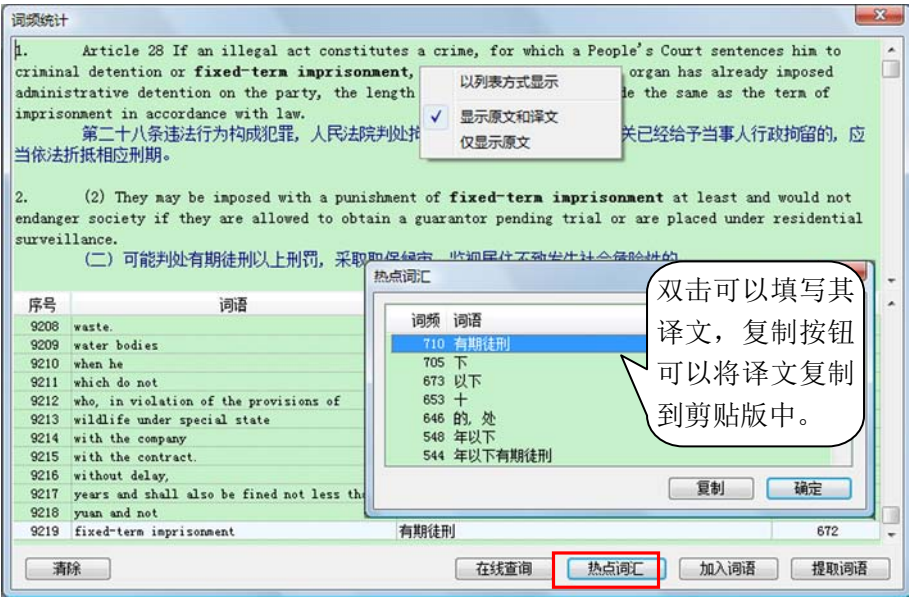
译文

2006年德国世界杯上首次颁发了佳新秀，获奖者为德国队的卢卡斯·波多尔斯基。

## 四、词频统计热点词汇

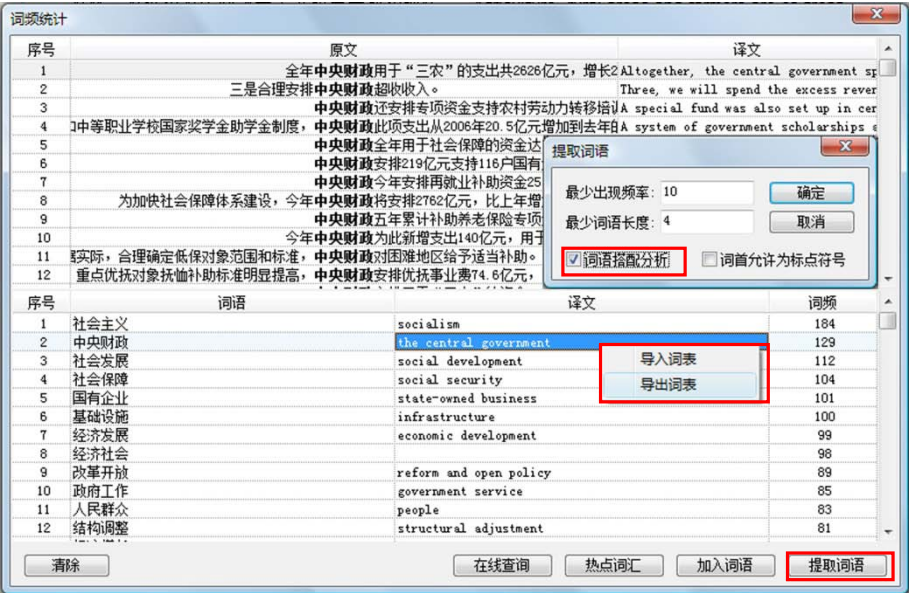
词频统计的结果，除了可以按表格方式显示外，还可以按“文本方式”显示，点击【右键菜单→以文本方式显示】功能可以切换到文本方式，如下图：在文本方式中，选择需要的词汇，可以利用【Ctrl+C】、【Ctrl+V】来复制粘贴。此外，文本方式显示可以设置为“仅显示原文”或“显示原文和译文”。

点击窗口下方的“热点词汇”按钮，用于计算双语文章中可能的译文。双击“热点词汇”窗口的词语，可以将它直接填写到译文区，点击“复制”按钮，可以把词语复制到剪贴板。



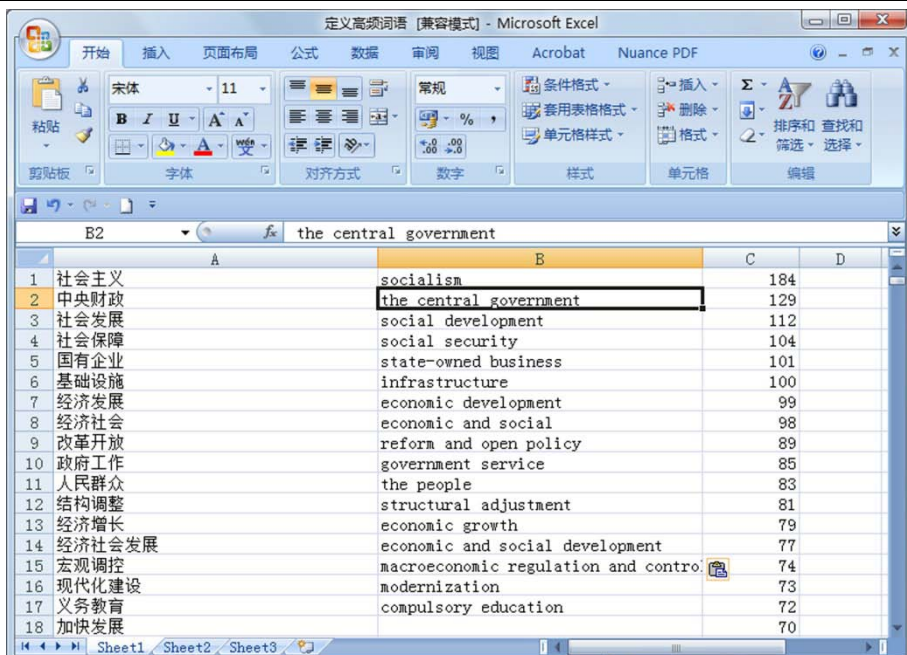
## 五、词频统计的词语搭配分析

为了进一步提高词频统计的准确性，在“提取词语”中加入“词语搭配分析”选项，使用该选项可以利用词语搭配的信息改进提取出来的词语，使词频统计的记录数减少一半左右，大大提高词频统计的实用性。例如：下图中，从“政府工作报告”文章中提取4字以上长度、出现频率10次以上的词语，不使用“词语搭配分析”选项的提取结果有6000多个，而使用“词语搭配分析”选项的提取结果是2000多个。将提取的词语排序后，会更加清晰（点击表头的词语列、译文列等可以排序），如下图。

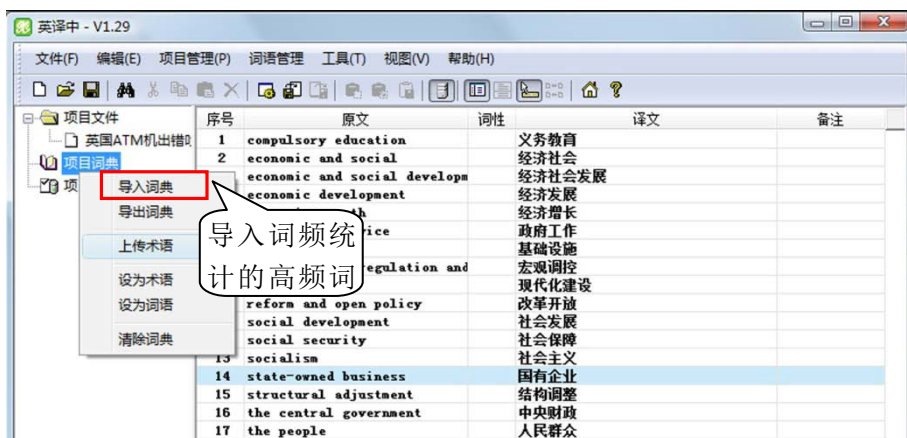


为了方便对词频统计的结果进行进一步的处理，可以将统计出的结果导出，或者导入已有的词表分析其在当前翻译项目中重复出现情况。例如：一个新的翻译项目，首先利用词频统计功能统计出项目中出现频率较高的词语，然后点击【右键菜单→导出词表】功能，将它导出成 EXCEL 文件（如下图），将这个 EXCEL 文件发给他入、或自己进行词语预先定义。





在 EXCEL 中将这高频词预先定义好后，就可以导入翻译项目的“项目词典”中，如下图所示。如果将它们设为术语，翻译时遇到这些词语时，就会自动提示在“术语区”中。



【右键菜单→导入词表】功能，可以检验自己定义的词表在当前翻译项目中使用情况，导入词表后，通过每个词显示的词频数值，就可以知道这些词语的重复出现情况了。对当前项目没有使用上的词语，其词频数值就显示为 0，点击“词频列”的表头，按出现次数排序就非常清楚了。

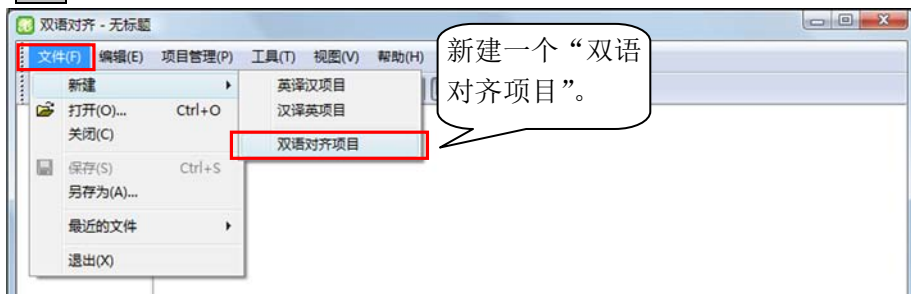


## 第十一章 双语对齐创建大型记忆库

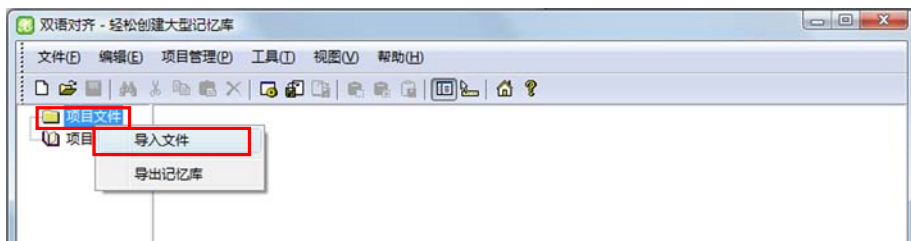
记忆库是译者的宝贵财富，拥有大型、高质量的记忆库，可以成倍地提高翻译的效率。记忆库的来源要么靠自己翻译的积累，要么靠收集。然而，自己翻译的积累，毕竟有限。现在网络资源丰富，双语资料随处可得，如何快速地将这些双语语料进行句级的对齐，创建大型的句库（记忆库）为译者服务？

雪人提供了一个功能强大的双语对齐工具，可以帮助译者快速创建大型的记忆库。**双语对齐项目中支持导入文件格式有：TXT、WORD、TMX、EXCEL、STM** 等等，导入 EXCEL 时会自动拆分句子进行对齐。其中 STM 格式需要标准版才能支持。下面具体讲述如何进行双语句级对齐？

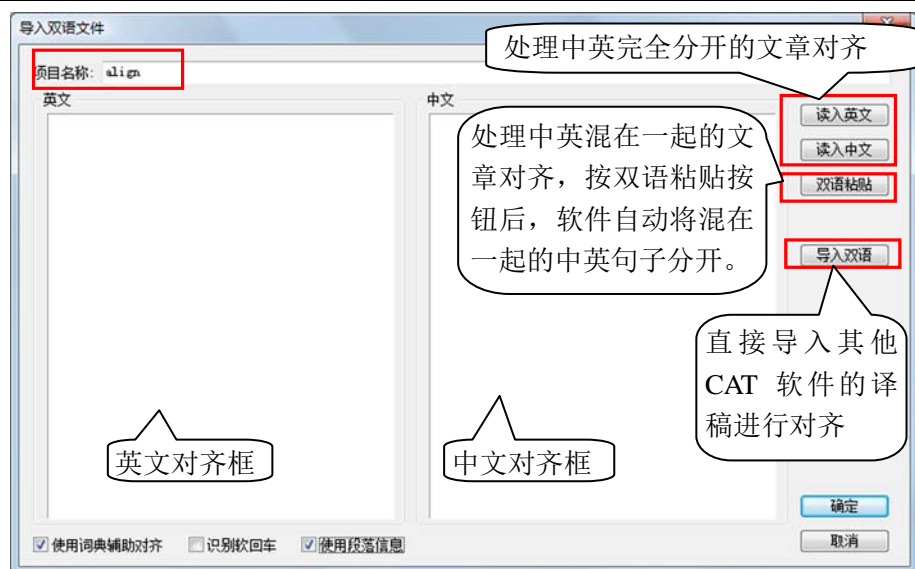
首先，点击【文件→新建→双语对齐项目】菜单，建立一个“双语对齐项目”，如下图：此时会弹出同新建翻译项目一样的“项目设置”对话框，用户可以不理会它，直接按**确定**按钮即可。



接着点击屏幕左侧的**项目文件**，然后点击【右键菜单→导入文件】功能，如下图：



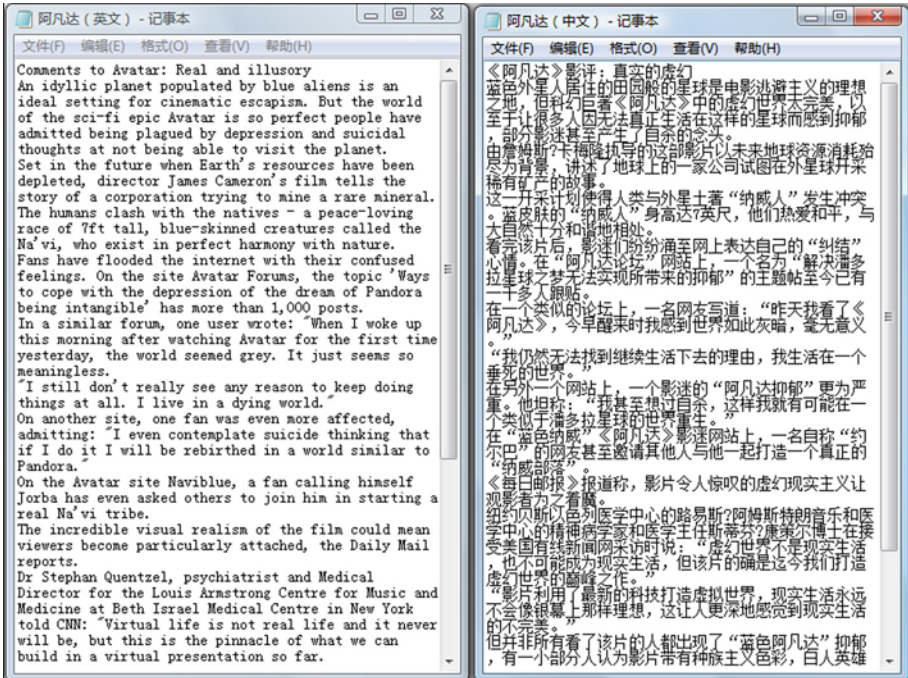
此时会弹出“导入双语文件”对话框，如下图：上方是项目名称，默认为“align”，用户可以修改，左侧是英文对齐框，右侧是中文对齐框，最右边是**读入英文**、**读入中文**、**双语粘贴**、**导入双语**等按钮。



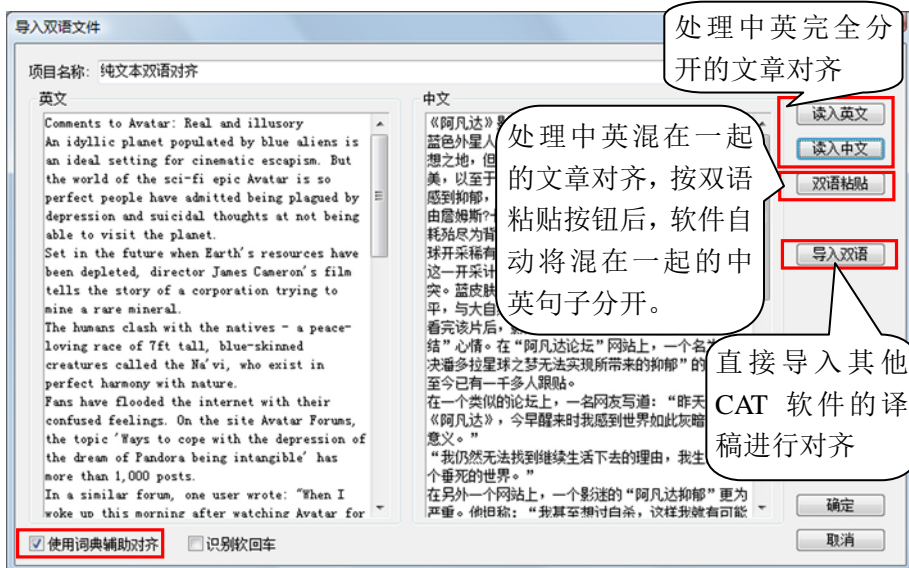
其中**读入英文**、**读入中文**按钮是用于处理中英完全分开的文章对齐，例如一篇中文、一篇英文，将这两篇文章进行一句一句的对齐。**双语粘贴**按钮是用于处理中英混在一起的文章对齐，按**双语粘贴**按钮后，软件自动将混在一起的中英句子分开。**导入双语**按钮是用于处理其他 CAT 软件的译稿，将在其他 CAT 软件中翻译过的译稿文件，直接导入进行句级对齐。

此外，窗口的下方还有“使用词典辅助对齐”、“识别软回车”、“使用段落信息”这三个选项，勾选“使用词典辅助对齐”和“使用段落信息”这两个选项，可以使对齐的准确性更高，建议用户“勾选”它们，“识别软回车”则根据具体的文档来确定，若你对齐的文档中含有大量的软回车，则就需要“勾选”这个选项了。

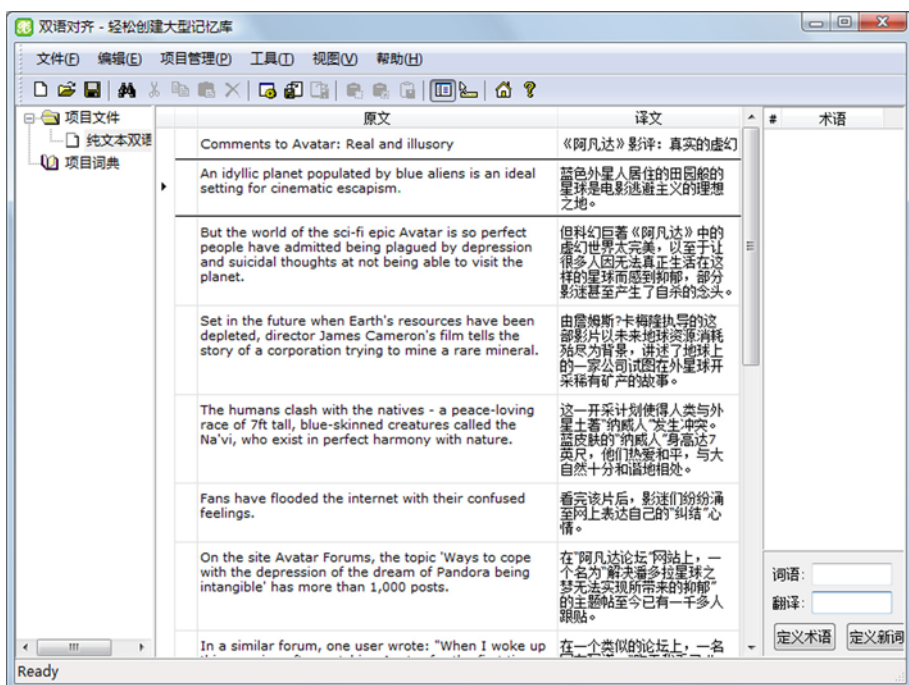
## 一、中英文分开的对齐



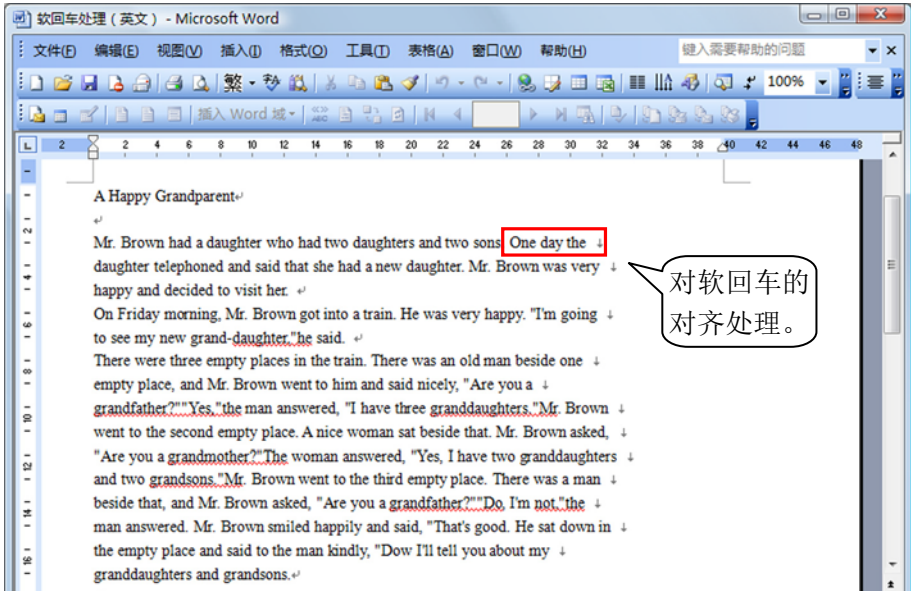
例如上图中，我们欲进行对齐的是两个 TXT 的文本文件，一个是英文，一个是中文，我们希望雪人能将它们一句一句地对应起来，做成句级对齐的记忆库。在下图的“导入双语”文件窗口中，我们点击**读入英文**按钮，读入一个英文的文件，英文显示在屏幕的左侧框中，另外点击**读入中文**按钮，读入一个中文的文件，中文显示在右侧框中。左下角的“**使用词典辅助对齐**”选项，默认是选择的，它可以提高对齐的准确度，但在对齐处理时会暂用一定的处理时间，比不选用它会稍微慢一些。读入后仍然可以在下图中编辑、修改，按**【Ctrl+A】**选择所有后按**【Delete】**可清除读入的内容。除了可以利用读入按钮读入文件，还可以利用快捷键，**【Ctrl+C】**、**【Ctrl+V】**直接将英文、中文粘贴到下图的对齐框中。



按**确定**按钮后, 自动对齐工作即刻完成, 如下图: 雪人已经将这两个文件中的一句中文、一句英文完整地对齐了, 而且准确率非常的高。

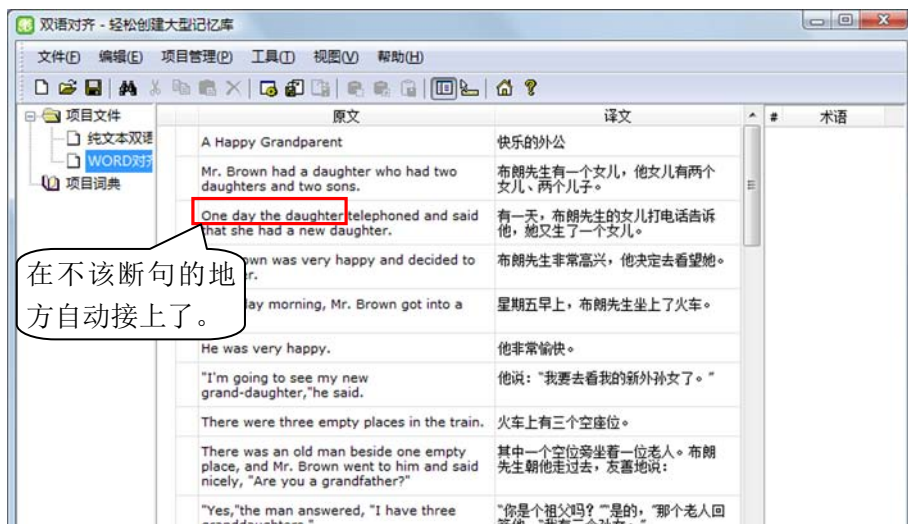
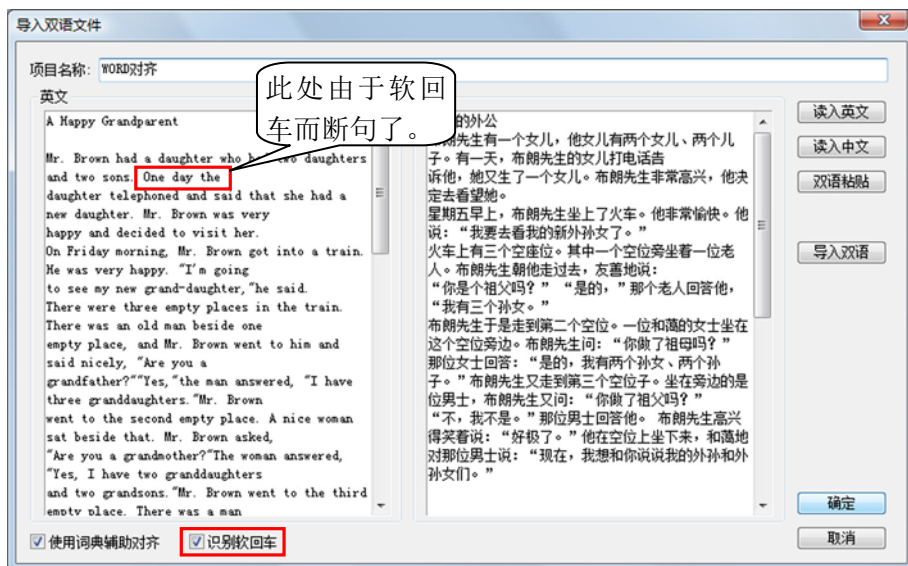


## 自动处理软回车



如上图中，有时我们会遇到一些文章中含有许许多多的软回车，它是由于排版等其他原因造成的，在不该断句的地方转行了，例如上图中，在“One day the”和“daughter”之间因为有个软回车而转行了，若不处理这些软回车，必然会影响对齐的效果。如果一个一个地手工删除软回车，会很耗时；雪人可以很好地自动处理软回车。我们将这两个 WORD 文档导入（**免费版都可以导入 WORD 文件来进行双语对齐的**），如下图：在屏幕的下方有一个“**识别软回车**”选项，勾选它后，按**确定**按钮。雪人已经自动处理了所有的软回车，在不该断句的地方自动接上了，而且对齐的效果很好！



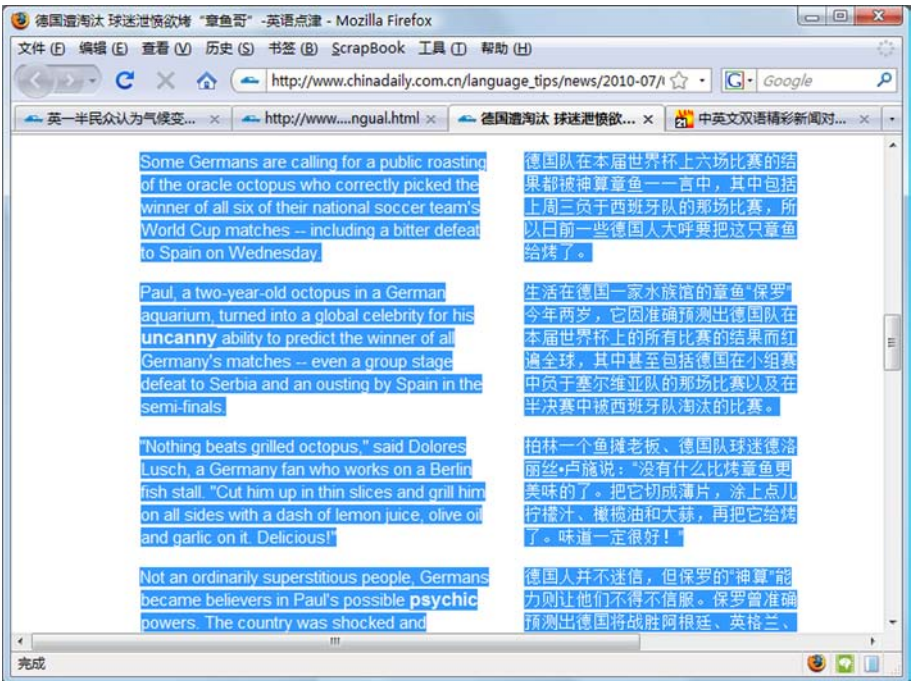
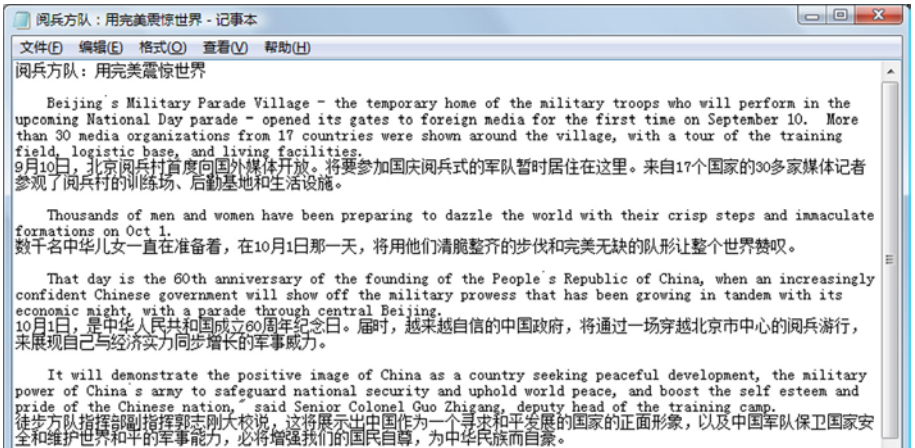


## 二、中英混排的对齐

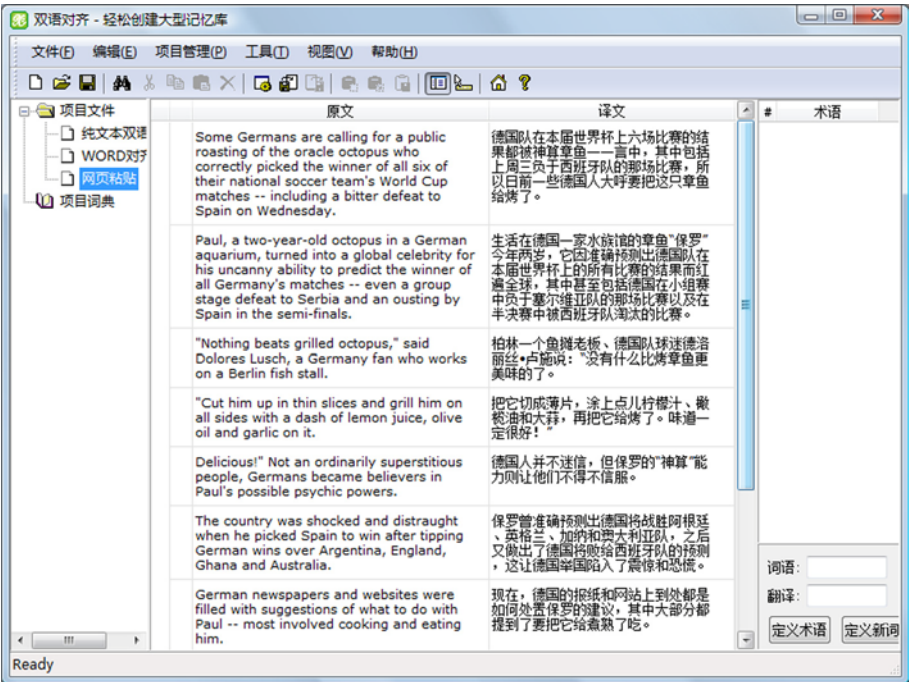
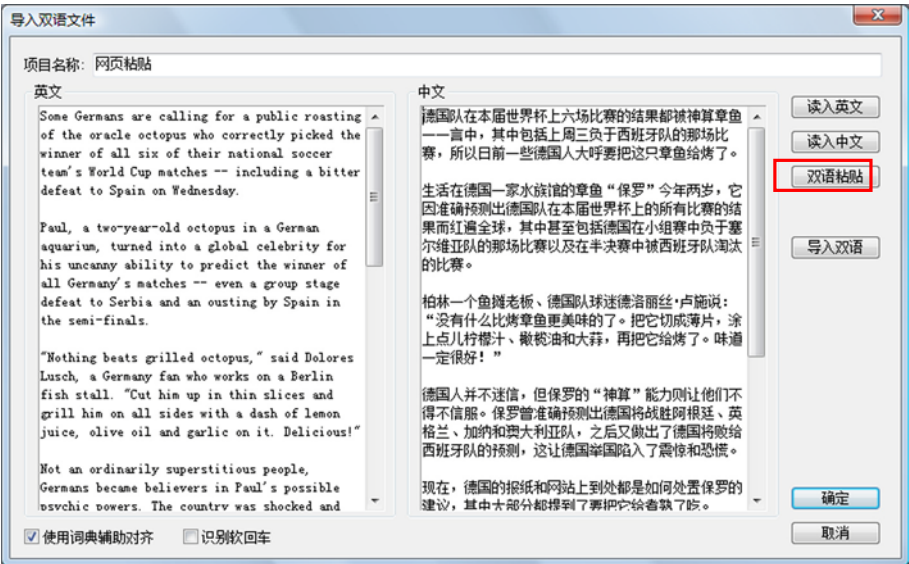
有时我们会遇到有些文章是中英混排的，它一段中文、一段英文，或者从网页中粘贴下来的，中英文混在一起，实在难以分成一个中文文件、一个英文文件。对处理这样文章，



有没有省时省力的方法呢？如下面的 2 个图：



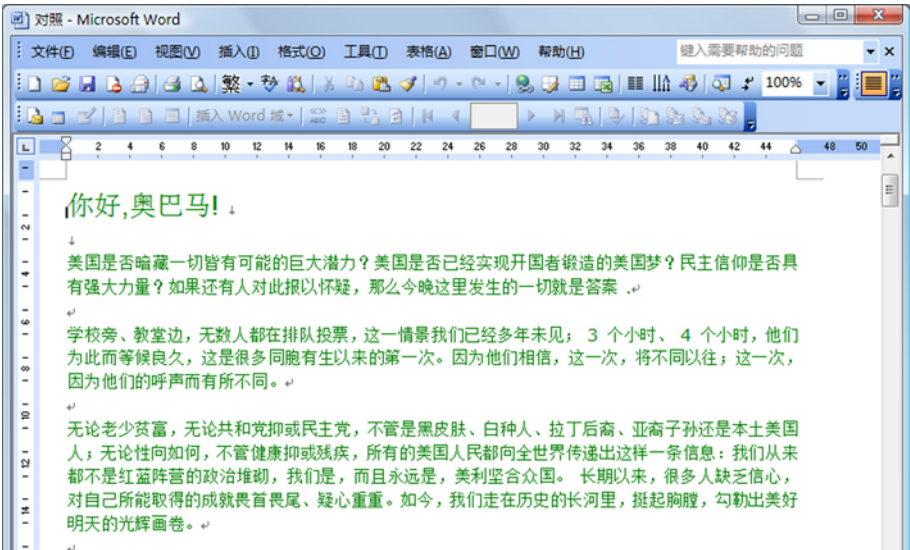
雪人可以很好地处理这种混排的对齐，以网页为例说明，我们一次选择了上图网页的中文和英文两个部分，按下【Ctrl+C】复制，然后在雪人中按**双语粘贴**按钮，如下图，此时雪人已经将刚才中英文混在一起复制的内容自动分开，分别粘贴在“英文对齐框”和“中文对齐框中”了。点击**确定**按钮，即可看到完美的一一对齐效果！



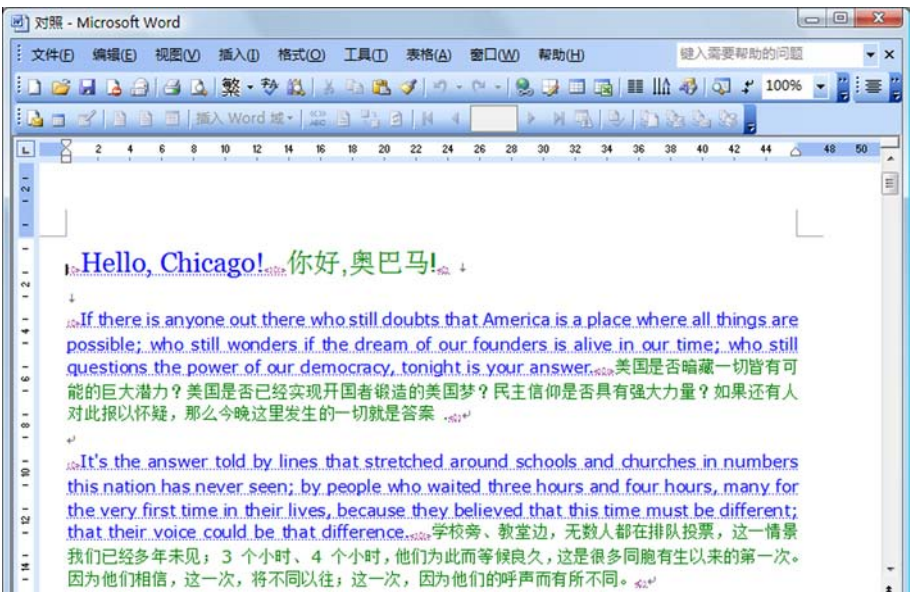
### 三、直接导入其他 CAT 软件译稿制作记忆库

如果你之前是用其他的 CAT 软件进行翻译的，雪人还可以直接导入其他 CAT 软件生产

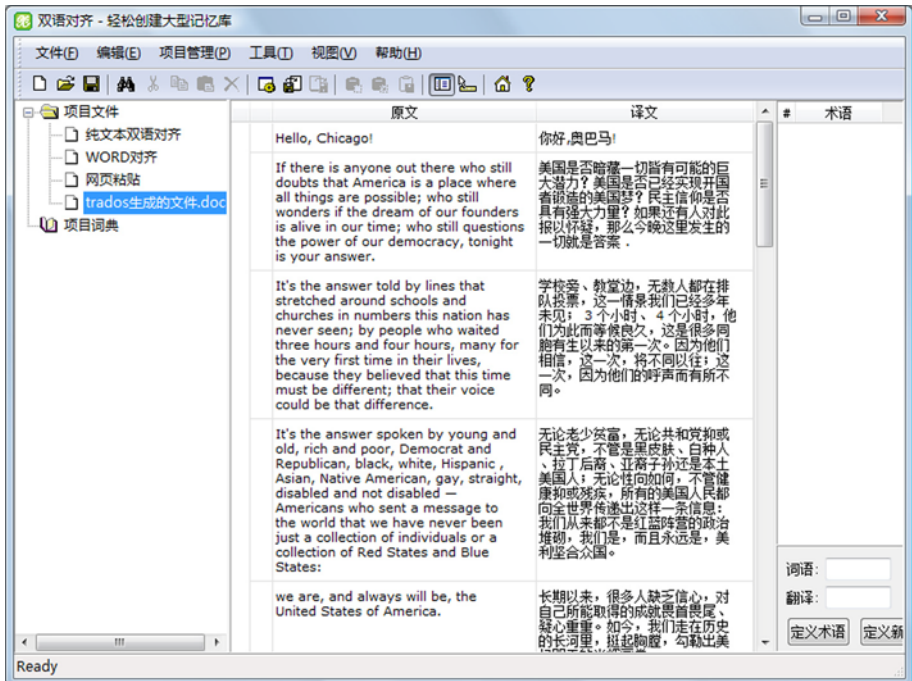
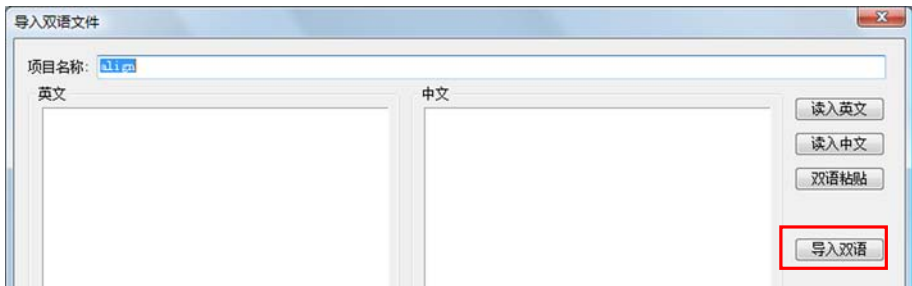
的译稿文件，将它们制作成一一对应的记忆库。目前可以直接导入的是 TRADOS 和雅信格式的。如下图，是用 TRADOS 生成的译文。



当使用 WORD 的“显示隐藏文件”功能时，可以看到其原稿如下图：



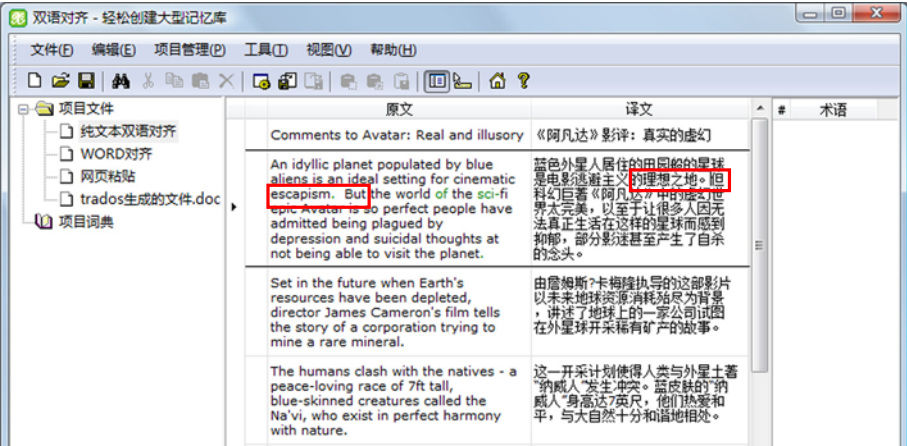
点击**导入双语**按钮，即可将 TRADOS 的译稿导入到雪人中，并自动对齐好，如下图：



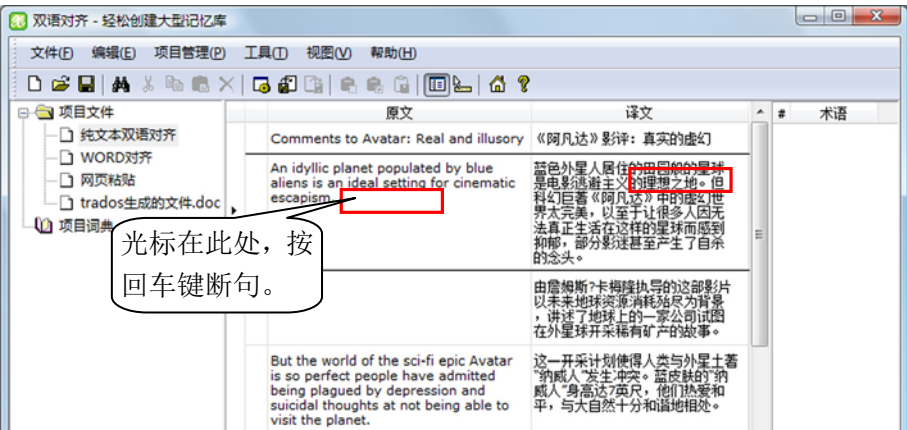


四、自动对齐后的编辑

虽然雪人的自动对齐效果不错，但也不是百分百准确，有时难免会出现不正确的情况。这时就需要用户手工修正了。雪人提供了非常方便的编辑功能，**自动对齐后的原文、译文都可以再修改、编辑**。如下图中，在“**But**”前本该断句，“**But**”后应该属于第二句。



我们将光标插入需要断句的地方，然后按**回车**键，即可在该处断开，如下图：这时我们看到原文、译文的移动是**不同步**的，即原文断开了，而译文并没有发生改变，同理，我们需要在译文相应的位置插入光标并按**回车**键，将译文断开。



当左右两边对齐后，我们按下**【F9】**键，将左右两句**锁定**，这时会出现一个“√”，表示这两句已经锁定，锁定后即可左右一起移动，不会再错位移动了。



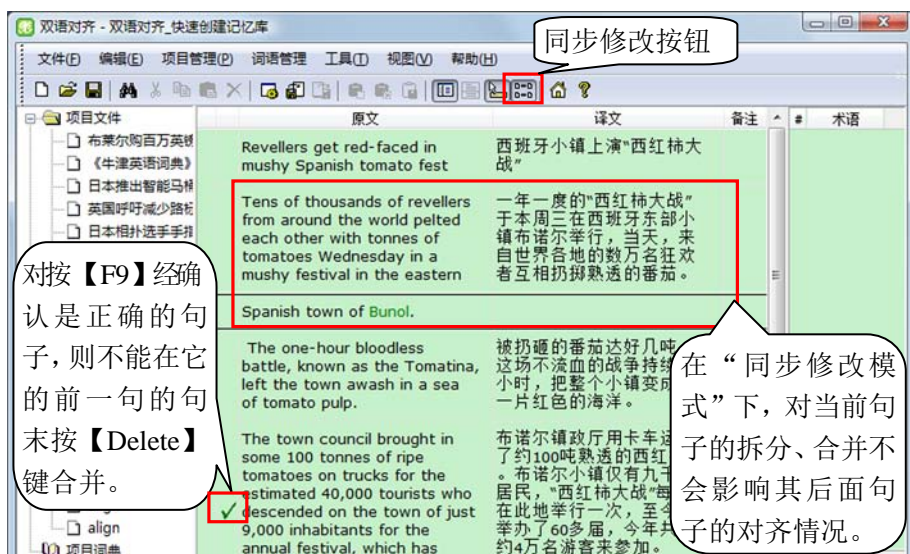
若发现原文、译文有错，可以对它们进行修改。若觉得某句的质量不高，日后不具备参考价值，可以按下【F2】键(【F2】为编辑、浏览切换键)，此时变成蓝色的长光标浏览状态，按【Delete】键可以整行删除；或者按【F4】键将出现“X”，表示该句不需要导出到记忆库中。

## 五、同步修改双语对齐

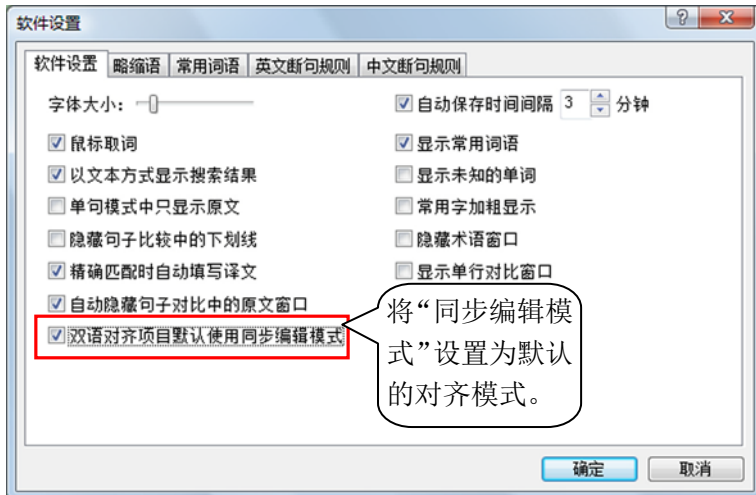
在双语对齐创建记忆库时，虽然大多数情况下能自动处理句与句的对齐，但有时仍会出现个别句子不能正确对齐的情况，此时就需要手工对齐了。有时出现不能正确对齐的原因是由于某个句子没有对齐好，而影响了后面的句子也没能正确对齐，只要将该句正确对齐了，其后面的句子也就自动对齐好了。所以软件默认状态下，如果对某个句子进行了拆分或合并的操作，其后面各句对的对齐会自动更改。

如果不希望对当前句子进行拆分、合并而影响其后面句子的对齐情况，可以点击工具栏的 “同步修改模式”按钮。再次点击该按钮则回复到默认的设置状态。



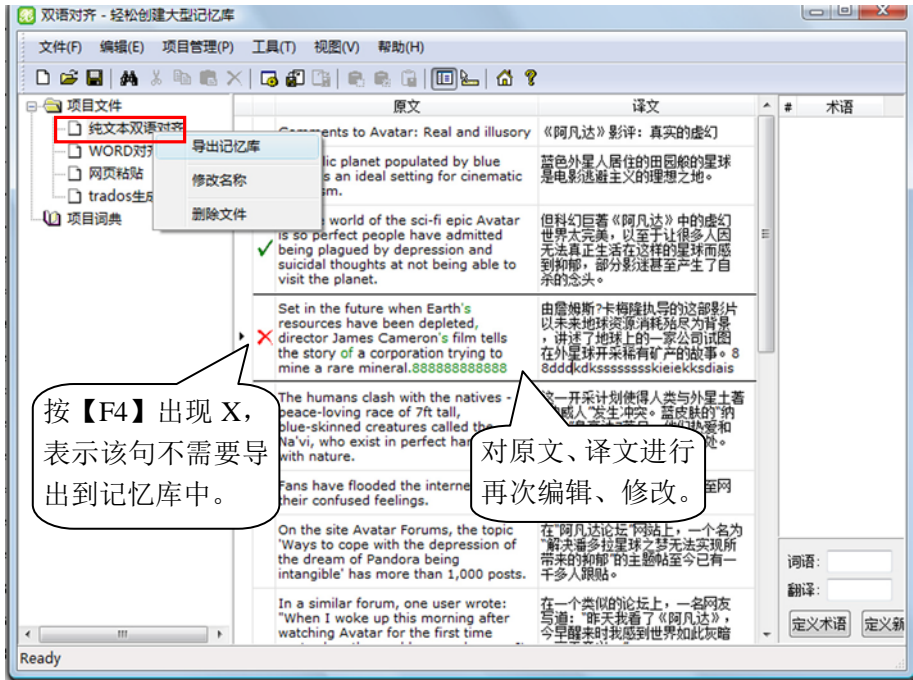


如果希望将“同步编辑模式”设置为默认的对齐模式, 可以点击【工具→软件设置】菜单, 在软件设置窗口中的“双语对齐项目默认使用同步编辑模式”选项打勾, 如下图: 修改设置后需要重新启动软件才能生效。



## 六、导出记忆库

修改好后，点击屏幕左侧的某个具体的文件，例如点击“纯文本双语对齐”这个文件，然后点击【右键菜单→导出记忆库】功能，可以将它**单独**导出记忆库。若点击左侧的**项目文件**文件夹后，再点击【右键菜单→导出记忆库】功能，则是将该文件夹下面的**所有文件**都导出到一个记忆库中。对齐后的内容可以导出为：STM(雪人格式的记忆库)、TMX (CAT 软件的记忆库标准格式)、EXCEL 格式。

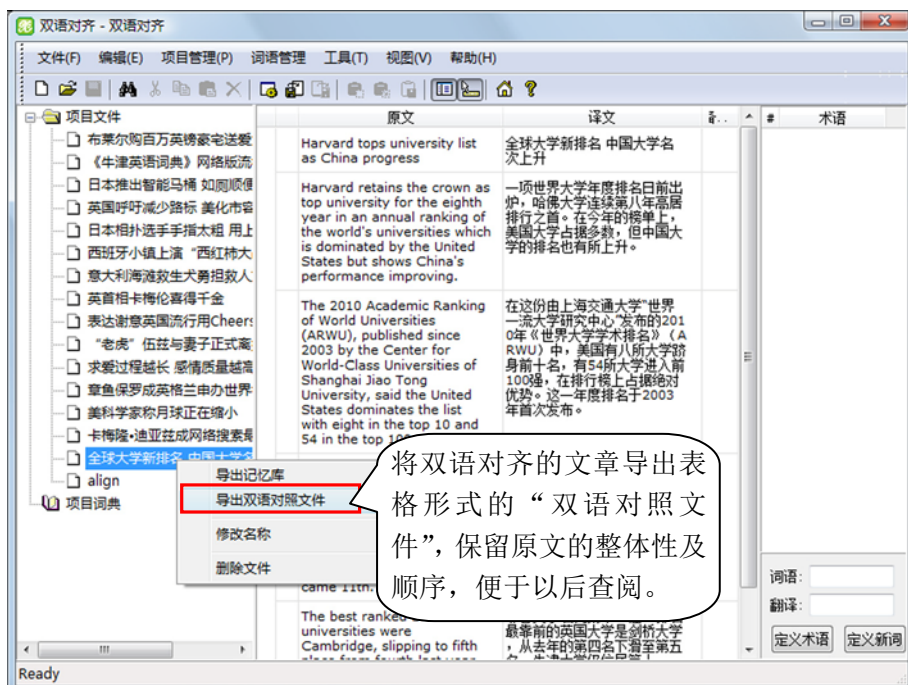


## 七、导出双语对照文件

用户自己收集、整理的文章利用雪人的双语对齐功能，进行句级对齐后便形成了记忆库，记忆库中的句子是独立的，没有了篇章的整体性。为了使这些资料日后更易阅读，雪人在“双语对齐”项目中都支持了将对齐的内容导出为“双语对照文件”，这样可以保留原文的顺序及篇章的整体性，方便查阅上下文！

点击左侧的具体文章，然后点击右键菜单的“导出双语对照”功能可以将它导出，如下图：导出的文件格式是 html 格式，可以直接用 IE 打开，或者用 EXCEL 打开，它虽然被

拆分成了一句一句，但仍然保留着原来的行文顺序，方便阅读（“导出双语对照”功能需要标准版才能支持）！



全球大学新排名 中国大学名次上升 - 由 Dell 提供的 Internet Explorer

D:\translation\雪人记忆库\全球大学新排名.HTM

Google

收藏夹 获取更多加载项 新建文件夹

全球大学新排名 中国大学名次上升

页面(P) 安全(S) 工具(O)

### 全球大学新排名 中国大学名次上升

原文	译文	备注
Harvard tops university list as China progress	全球大学新排名 中国大学名次上升	
Harvard retains the crown as top university for the eighth year in an annual ranking of the world's universities which is dominated by the United States but shows China's performance improving.	一项世界大学年度排名日前出炉，哈佛大学连续第八年高居排行之首。在今天的榜单上，美国大学占据多数，但中国大学的排名也有所上升。	
The 2010 Academic Ranking of World Universities (ARWU), published since 2003 by the Center for World-Class Universities of Shanghai Jiao Tong University, said the United States dominates the list with eight in the top 10 and 54 in the top 100.	在这份由上海交通大学“世界一流大学研究中心”发布的2010年《世界大学学术排名》(ARWU)中，美国有八所大学跻身前十名，有54所大学进入前100强，在排行榜上占据绝对优势。这一年度排名于2003年首次发布。	
Joining Harvard in the top 10 were the University of California, Berkeley; Stanford; Massachusetts Institute of Technology; California Institute of Technology; Princeton; Columbia and Chicago. Yale came 11th.	除哈佛大学外，跻身前十名的美国大学还包括：加州大学伯克利分校、斯坦福大学、麻省理工学院、加州理工学院、普林斯顿大学、哥伦比亚大学和芝加哥大学。	
The best ranked British universities were Cambridge, slipping to fifth place from fourth last year, and Oxford retaining the 10th position.	耶鲁大学名列第十一位。排名学是剑桥大学，从去年的第四名下滑到第十位，牛津大学仍位居第十。	
Overall the number of British universities in the top 500 rankings dropped to 38 from 40.	总体来看，排在前500名的英国大学数量从40所降到了38所。	

完成 计算机 | 保护模式: 禁用 100%

用 IE 打开导出的双语对照文件，保留原文的行文顺序。

Microsoft Excel - 全球大学新排名

文件(F) 编辑(E) 视图(V) 插入(I) 格式(O) 工具(T) 数据(D) 窗口(W) 帮助(H) Adobe PDF(B)

100%

A1 全球大学新排名 中国大学名次上升

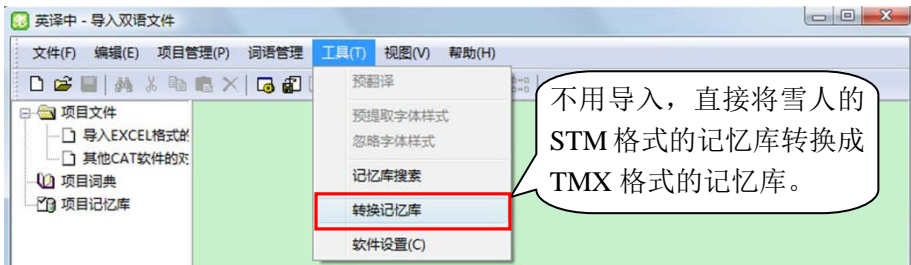
	A	B	C	D
1	全球大学新排名 中国大学名次上升			
2	原文	译文	备注	
3	Harvard tops university list as China progress	全球大学新排名 中国大学名次上升		
4	Harvard retains the crown as top university for the eighth year in an annual ranking of the world's	一项世界大学年度排名日前出炉，哈佛大学连续第八年高居排行之首。在今天的榜单		
5	The 2010 Academic Ranking of World Universities (ARWU), published since 2003 by the Center for World-Class Universities of Shanghai Jiao Tong	在这份由上海交通大学“世界一流大学研究中心”发布的2010年《世界大学学术排名》(ARWU)中，美国有八所大学跻身前十		
6	Joining Harvard in the top 10 were the University of California, Berkeley; Stanford; Massachusetts Institute of Technology; California Institute of	除哈佛大学外，跻身前十名的美国大学还包括：加州大学伯克利分校、斯坦福大学、麻省理工学院、加州理工学院、普林斯顿		
7	The best ranked British universities were Cambridge, slipping to fifth place from fourth last year, and Oxford	耶鲁大学名列第十一位。排名最靠前的大学是剑桥大学，从去年的第四名下滑		
8	Overall the number of British universities in the top 500 rankings dropped to 38 from 40.	总体来看，排在前500名的英国大学数所降到了38所。		
9	But the ranking, initially set up to find the global standing of Chinese universities, showed Asian universities were advancing up the list with 106 from	但从今年的排名可以看出，亚洲大学有所上升，亚太地区共有106所学校进500强，中国大学的排名也有所提升。		
10	"While the ranking methodology has been kept the same, the number of top 500 Chinese universities reaches 34 in 2010, which is more than double that in 2004 (16)," Shanghai Jiao Tong University said in a statement.	上海交通大学在一份声明中说：“在排名方法不变的情况下，今年进入全球500强的中国大学有34所，比2004年(16所)增加了一倍多。”		

全球大学新排名 /

用 EXCEL 打开导出的双语对照文件，保留原文的行文顺序。

## 八、记忆库的直接转换

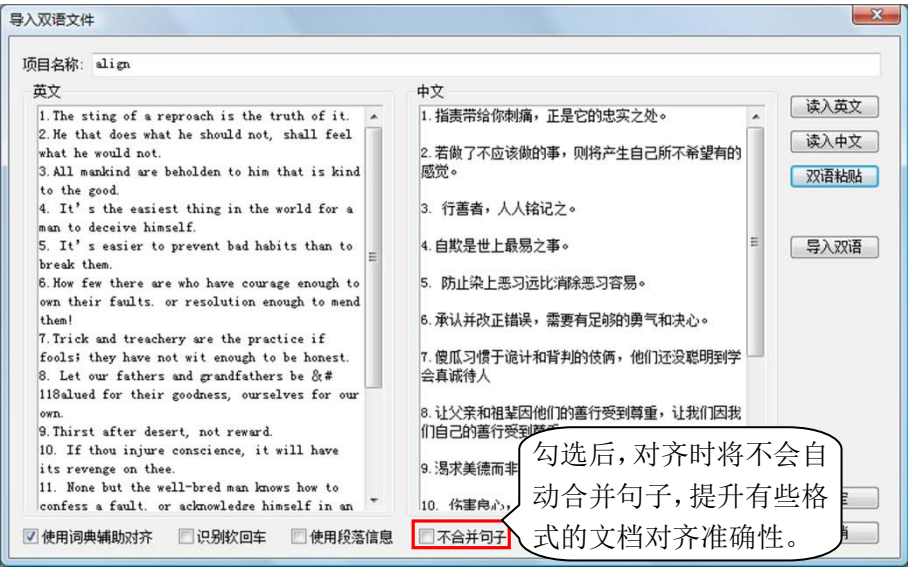
STM 是利用雪人 CAT 生成的记忆库文件格式，如果需要转换成通用的记忆库格式 TMX，除了通过导入 STM 记忆库后再导出 TMX 记忆库外，在 V1.23 版中，增加了“记忆库直接转换”功能，可以直接进行转换。对于转换大型记忆库，可以节省导入的时间和避免记忆库太大无法导入的问题。在标准版中点击【**工具菜单**→**转换记忆库**】功能，如下图：然后双击需要转换的 STM 格式的记忆库，在弹出的窗口中输入新的 TMX 格式记忆库的文件名即可快速转换了。



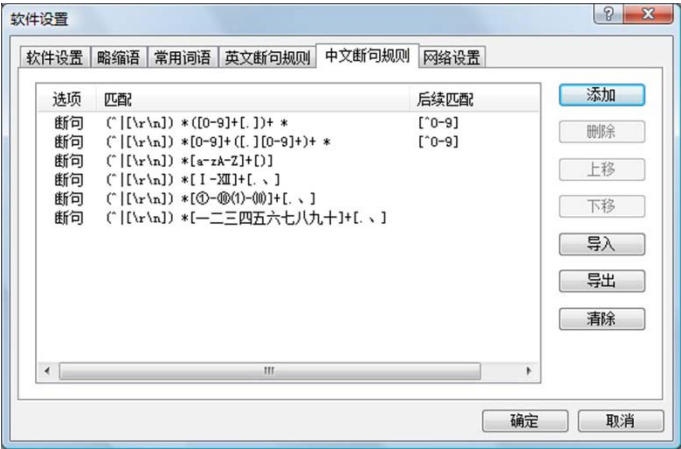
## 九、双语对齐的不合并句子选项

V1.29 版在双语对齐中加入“**不合并句子**”的选项，可以避免因为合并句子而造成的对齐问题。例如：下图中需要对齐的两篇文档中，含有很多的序号。做成记忆库的例句当然不希望还带着序号，但手工一个一个地删除这些序号会比较麻烦。我们希望各句中、英文的序列号处能自动断开，导出记忆库时忽略所有的序号。



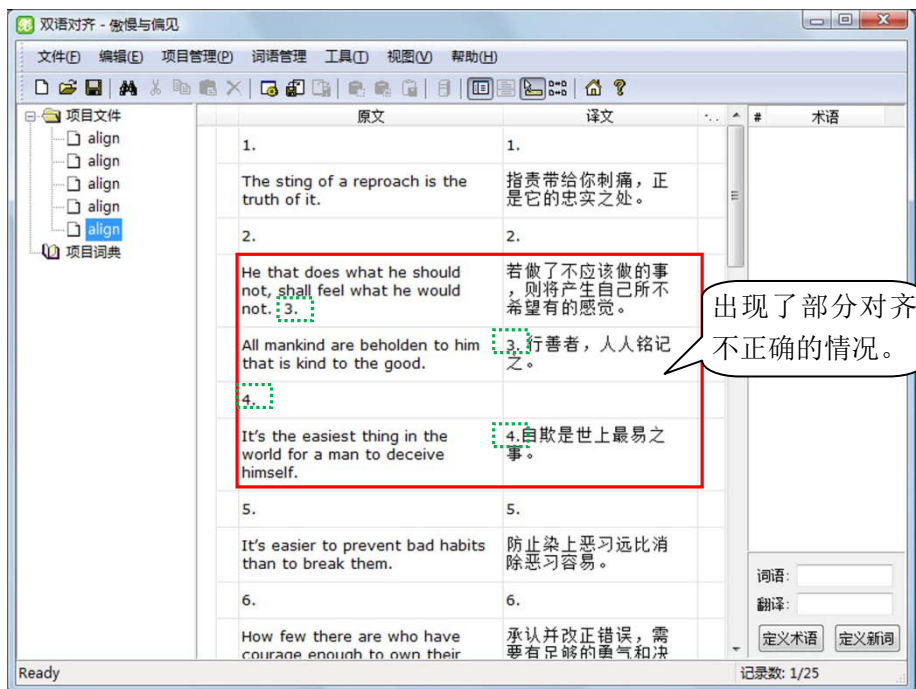


为此, 我们可以点击【工具→软件设置】中定义断句规则, 将序号单独断句。例如: 下图中设置的断句规则, 就可以处理 1. XXX 2. XXX, 1、XXX 2、XXX, a. XXX b. XXX, ①、XXX ②、XXX, I. XXX II. XXX, 一、XXX 二、XXX 等等情况下的序号单独断句。(这个断句规则文件放在免费版的“断句规则样例”文件夹中, 用户可以导入直接使用)。

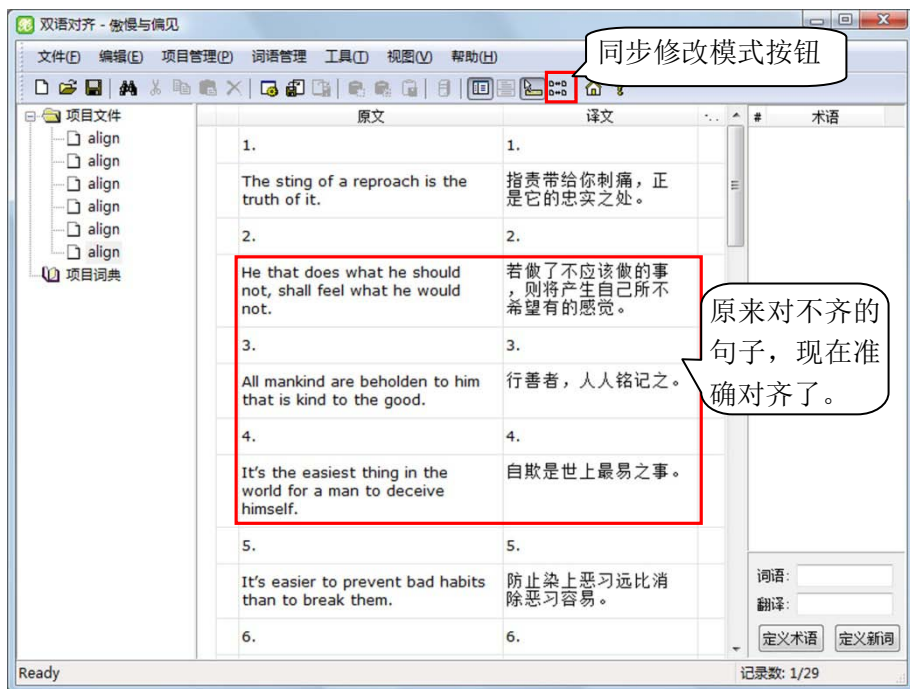


设定好断句规则后, 我们先不勾选“**不合并句子**”这个选项, 来看看对齐的效果, 对齐结果如下图: 通过断句规则, 序号基本上单独断句并对齐; 但第 3、第 4 句出现了对齐不正确的情况, 发生这种情况是因为软件在对齐时进行了句子的合并。





如果我们勾选“**不合并句子**”这个选项，再次重新对齐，则对齐的结果就不同了，原来第3、第4句已经准确对齐了。效果如下图：所有的序号都被单独断成了一句，导出记忆库时，这些序号行会被自动忽略，做成记忆库的句子将不会出现序号。



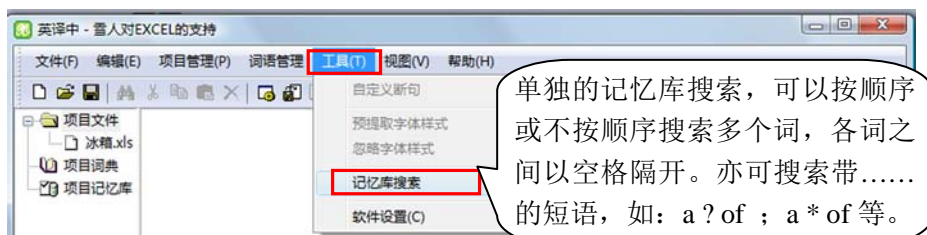
使用“**不合并句子**”选项方式进行的对齐，如果需要进行修改，建议点击屏幕工具栏的“**同步模式**”按钮，在“同步模式”中维护起来会更加方便。

“**不合并句子**”选项对句子一一对应的文章的对齐效果会比较好。对那些一句原文对应多句译文的情况，对齐时由于限制自动合并句子，则对齐效果反而还会差一些。用户可以两种方式都测试一下，哪种对齐效果好就选用哪种。

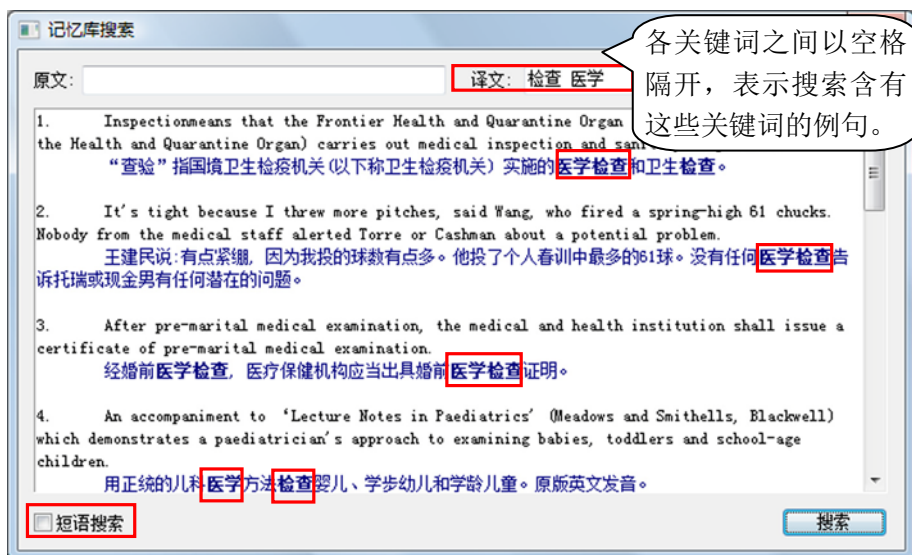
## 第十二章 独立的记忆库搜索

新版本中增加了对“记忆库”的独立搜索功能，搜索条件可以自由设定，搜索的范围包括当前的“项目记忆库”和“添加引用”的外部记忆库。

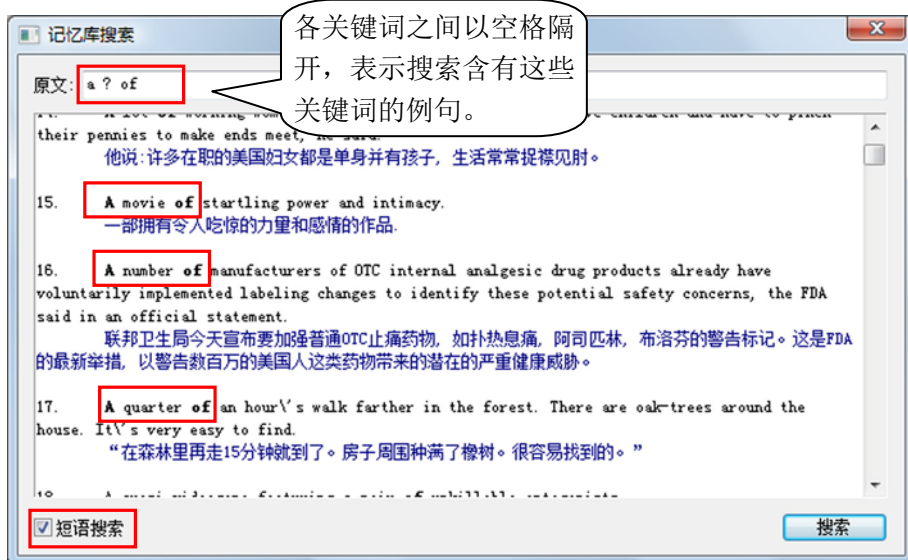
点击【工具→记忆库搜索】菜单，弹出窗口如下图：



搜索条件灵活多样，可单独搜索原文或单独搜索译文，当原文和译文输入框中同时输入内容是，表示搜索的结果需要同时满足条件才显示出来。例如：下图中，在译文搜索框中输入了“检查 医学”这两个关键词，中间并以空间隔开，则表示搜索译文中含有“检查 医学”这两个关键词的例句，注意：在窗口的左下角有一个“短语搜索”选项，默认是不打勾的，则表示不限定这两个关键词的出现顺序，只有例句中包含它们就会被搜查出来。若将“短语搜索”选项打勾，则表示搜索短语，各关键词之间是有序的，只有符合各关键词的输入顺序的例子才会显示出来。



短语搜索：点击搜索窗口左下角的“短语搜索”选项，将它设为“√”。在短语搜索中，有两个常用的特殊符号：“?”和“\*”，“?”表示匹配一个词，“\*”表示匹配任意个词。例如在下图的原文搜索框中输入了“a ? of”则表示搜索 a 后面仅跟一个词的短语，其结果如下：有 A movie of、A number of 等等。若将搜索条件设为“a \* of”，则表示搜索 a 后面任意字符，然后接 of 的句子了。



若在译文中亦同时输入搜索条件，则表示同时搜索满足这两个条件的例句。例如下图中：原文的搜索条件是“a ? of”，译文的搜索条件是“双倍”，其从记忆库中搜索到的结果是“a couple of”。



使用“?”和“\*”进行中文的搜索时，注意不要留空格，因为空格也是作为一个匹配单位来计算的。

## 第十三章 自定义语法规则翻译

无论是英语还是汉语，都有一定的语法规则可循。为了使翻译的结果更加准确、高效率，雪人提供了强大的“**用户自定义语法规则翻译**”功能，用户可以在规则词典中预先将这些规律定义成语法规则，这样就能进一步提高取词和翻译的准确性。例如“be satisfied with…，对…感到满意”，若将它定义到规则词典中，则以后若遇到“be satisfied with you”就可以准确翻译出：“对**你**感到满意”，若遇到“be satisfied the result”就可以将它准确译成：“对**这个结果**感到满意”。用户可以不断添加新的语法翻译规则，而且亦可与其他用户共享、共建语法翻译规则库。

常用的句型或语法规则，都是以某些单词或某类单词为核心，单词之间的搭配是固定的，在造句时要替换的只是其中的某些单词、短语或词组。我们可对这些规则进行归纳、总结，固定部分的单词和释义不变，替换的部分用“\*”表示，与其对应的译文用“#”表示。下面将详细介绍如何进行语法规则的定义：

### 一、什么是类别

为了使定义的规则更加灵活和实用，我们支持用户自定义“**类别**”属性，把单词或词组进行归类，定义它们的“属性”，单词和词组的词性及属性在系统中统称为“**类别**”，“类别”的名称用户自由定义。**通过“类别”的概念，使所有定义的语法模式有了继承性。**例如：我们把“Man”、“Girl”、“Boy”、“Women”等单词的属性定义为“人”，同时定义类“人”的单位（量词）为“个”，当这些单词与冠词(a, the, this, that, these, those)、数词(one, two, 1, 2, …)等搭配时，应加上量词“个”。为此。我们定义：“冠词”+“人”为一个模式，“数词”+“人”为另一个模式。这样，系统中所有属性为“人”的单词或词组将继承该模式。**“类别”放在“{}”中，其释义位置用“@”表示，翻译时用“{}”对应单词的释义替换模式中的“@”位置。**

### 二、规则定义的常用符号：

\*——用于对**原文**的定义，表示匹配当前位置的一个词或一个短语。（注意：\*不能出现在句首）

#——用于对**译文**的定义，表示对应于原文\*号位置的具体译文，当原文出现多个\*号时，则#1表示第一个\*号的内容，#2表示第二个\*号位置的内容，依次类推。

{ }——用于对**原文**的定义，表示引用“**类别**”，用户可自行定义，“类别”名称可以是中文名称或英文名称，另外软件有一些已经默认定义好的类别，用户可直接使用，软件中预定义类别有：

{digits}——表示数字类别

{label} ——表示数字与字母组合类别，如 a1；

{词性}——表示词性，其中 {n} 指名词，{ns} 地名，{nr} 人名，{v} 动词，{adj} 形容词，{adv} 副词，{pron} 代词，{int} 叹词。

除此之外，**用户可以自定义任意的“类别”名称，引用时只要用“{}”括住所定义类别名称就行了**。如：{xxxx}。

@ ——用于对**译文**的定义，表示对“类别”的翻译。@1 表示翻译原文“第一个{}”类别的内容，@2 表示翻译原文“第二个{}”类别的内容，依次类推。

^——表示以句首开始。

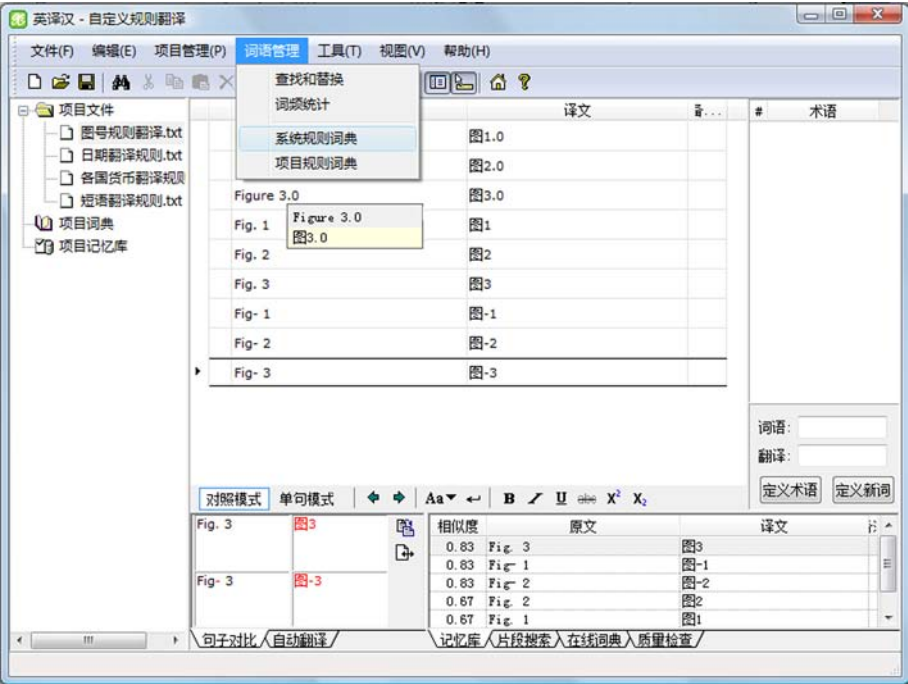
\——转义字符，由于软件对“\* {} @ # ^”等这些字符有特别的作用，所以若真正需要输入这些已经被定义的符号时，就需要在它的前面加上“\”，例如：\\*表示输入\*这个符号， \@表示输入@这个符号。

**注意：规则中使用的“\*”、“#”、“{}”和编号“1”、“2”...9 均为半角英文字符。单词之间、单词和符号之间要留一个空格（半角）。如果译文为空，则该条规则会被忽略。另外，错误的规则会以红色显示。**

语法规则翻译词典分为：“系统规则词典”和“项目规则词典”两个翻译规则库，其中在“系统规则词典”中定义的规则，可以应用于所有的翻译项目中，而在“项目规则词典”中定义的规则，则仅适用于当前的翻译项目中。“系统规则词典”保存在 SCAT 的安装目录下，文件名为“rule.dbf”；“项目规格词典”则保存在项目文件中。在“规则词典”窗口点击右键菜单，可以导入、导出和删除规则，按住【Ctrl】或【Shift】可以选择多个规则进行删除。

点击【**词语管理→系统规则词典或项目词典规则**】可以打开规则定义窗口，如下图：我们将通过几个具体的实例应用，来讲解怎样根据需要自定义翻译规则。





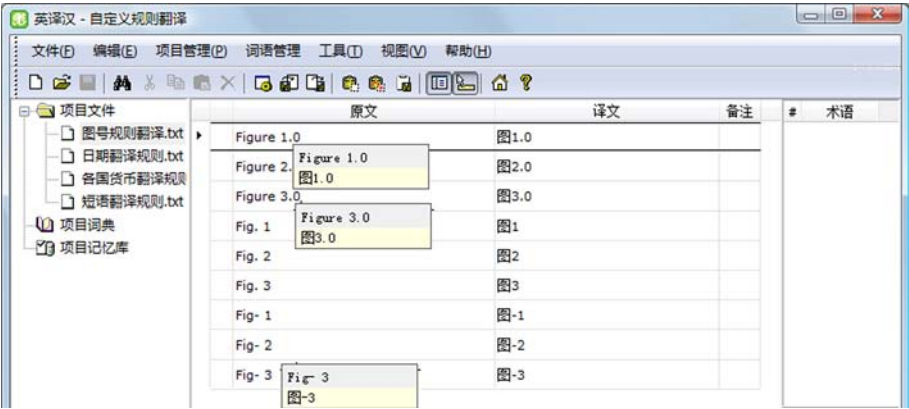
例 1：图号规则翻译。在翻译时，经常会遇到不少插图，而且每个图有序号。表达方式有些文章是用 Figure 1.0、2.0、3.0 表示，有些是用 Fig-1、-2、-3 等等，例如上图图中所示。我们希望通过定义翻译规则，日后不管遇到那种表示方法，都可以自动将它翻译为：“图 1.0、图 2.0 或图-1、图-2 等。

由于我们希望所定义的翻译规则在以后所有的翻译项目中都可以使用，所以在此点击【词汇管理→系统规则词典】菜单，弹出“系统规则词典”窗口，如下图：我们在“原文列”中输入了“Fig- {digits}”，其中“Fig-”是英文中一种常用的图号表达方式，而“{digits}”是软件默认类别，表示数字；在“译文列”中，输入了“图-@1”，其中“@1”表示在这个位置上翻译“{digits}”的内容，由于原文{digits}表示数字，则在翻译时“图-@1”就变成：图-1、图-2、图-3 等等了。

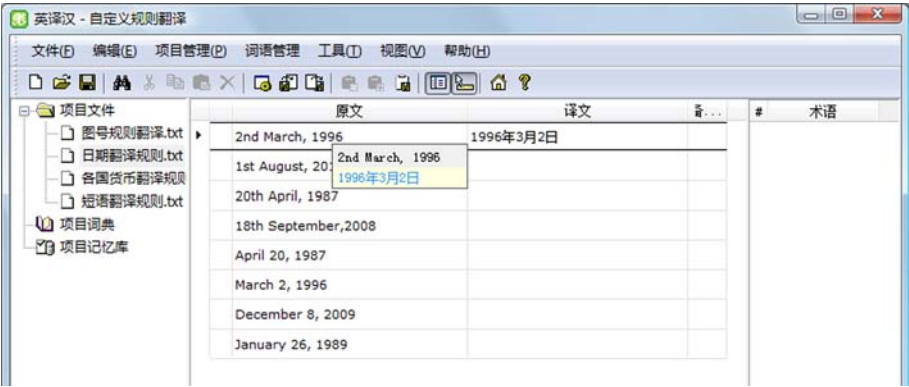


定义好规则后，退出“系统规则词典”窗口，当鼠标分别指向 Figure 1.0、2.0、3.0

时，屏幕取词窗口就会自动将它们翻译成“图 1.0、图 2.0、图 3.0”，当鼠标分别指向 Fig-1、-2、-3 时，就会自动将它们翻译成“图-1、图-2、图-3”，如下图：



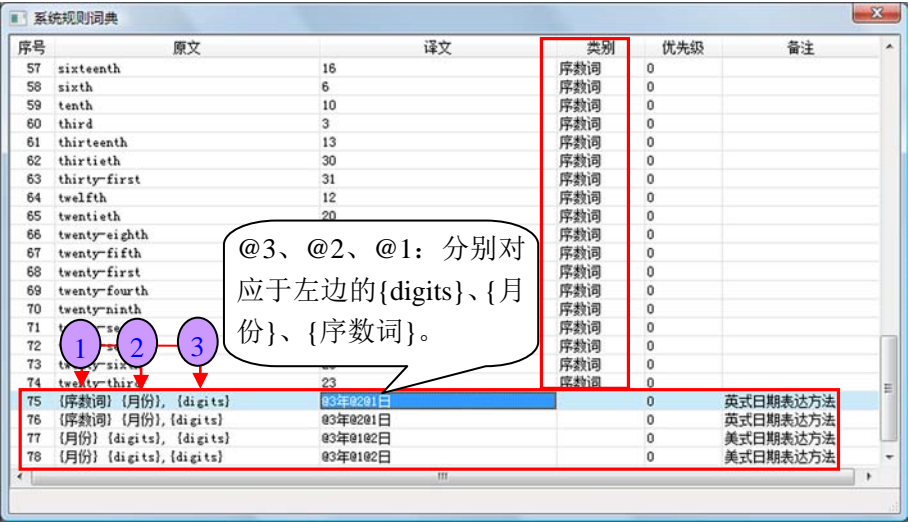
例 2：日期规则翻译。英语日期表达方法有两种：英式日期表达方式是序数词（日）、月份、年；美式日期表达方式是：月份、日期、年。如下图：我们希望通过简单的翻译规则定义，能正确翻译出所有的日期。



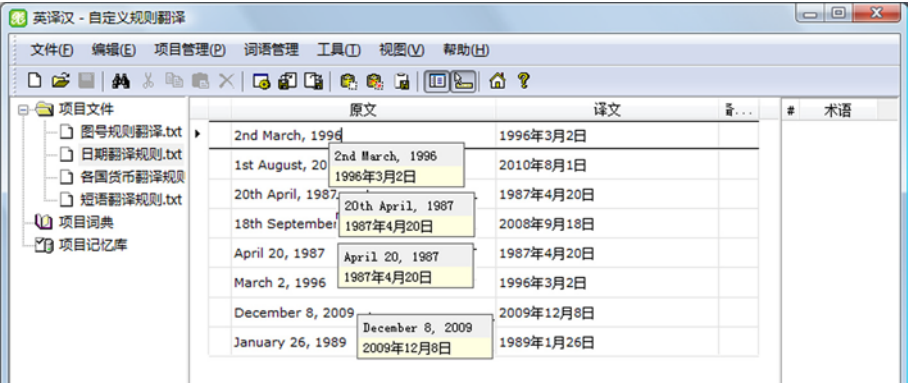
点击【词语管理→系统规则词典】菜单，在“系统规则词典”窗口中将各个月份及其缩写的原文、译文定义好，并在“类别”列中输入“月份”二字，如下图。

序号	原文	译文	类别	优先级	备注
22	April	4月	月份	0	
23	Aug.	8月	月份	0	
24	August	8月	月份	0	
25	Dec.	12月	月份	0	
26	December	12月	月份	0	
27	Feb.	2月	月份	0	
28	February	2月	月份	0	
29	Jan.	1月	月份	0	
30	January	1月	月份	0	
31	Jul.	7月	月份	0	
32	July	7月	月份	0	
33	Jun.	6月	月份	0	
34	June	6月	月份	0	
35	Mar.	3月	月份	0	
36	March	3月	月份	0	
37	May	5月	月份	0	
38	Nov.	11月	月份	0	
39	November	11月	月份	0	
40	Oct.	10月	月份	0	
41	October	10月	月份	0	
42	Sept.	9月	月份	0	
43	September	9月	月份	0	
44	eighteenth	18	序数词	0	

然后输入 1-31 的序数词及缩写的原文、译文，并在“类别”列中统一定义它们为“序数词”。接着在“原文”列中输入了“{序数词}{月份}，{digits}”，其中“{序数词}”表示引用“类别”为序数词的内容，{月份}表示引用“类别”为“月份”的内容，{digits}表示数字类别；在“译文”列中输入“@3 年@2@1 日”，@3 表示翻译位置为 3 号的类别所对应的内容，@2 表示翻译位置为 2 号的“类别”所对应的内容，@1 表示翻译位置为 1 号的类别所对应的内容，如下图：另外有些日期的表达方式不是很规范，在月份与年之间没有加空格，有无空格在软件中会认为是不同的两个规则，所以，为了识别那些没有空格的日期，我们又增加了第二个规则“{序数词}{月份}，{digits}”，与第一个不同之处在于“{月份}，{digits}”之间没有空格。这样以后不管遇到有无空格的情况，都可以正确翻译了。



定义好日期翻译的规则后，退出“系统规则词典”窗口，如下图：当鼠标分别指向各种不同的时期时，屏幕取词窗口就会自动对它们做正确翻译了。



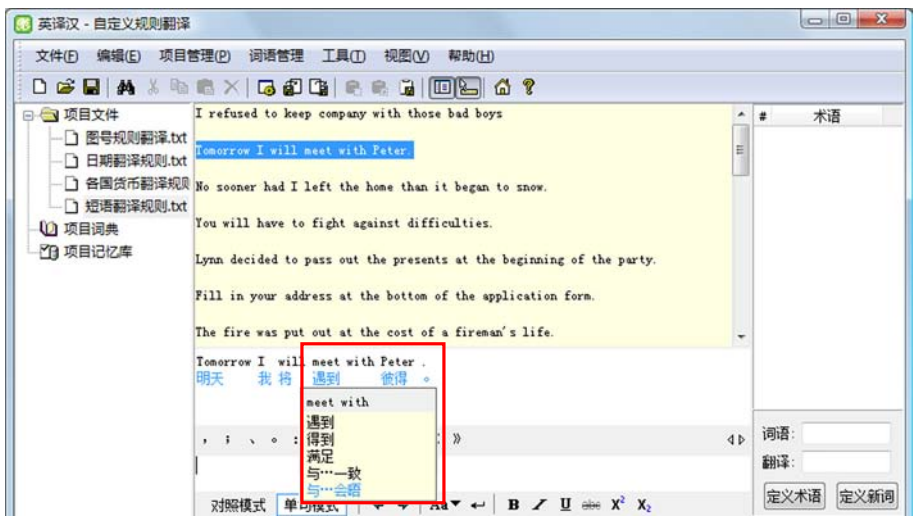
例 3：各国货币翻译。下图中在“系统规则词典”中定义了各种货币缩写及其对应的译文，如：RMB {digits}——@1 元人民币，USD {digits}——@1 美元，其中@1 表示第一个“{}”的内容。

序号	原文	译文	类别	优先级	备注
1	\${digits}	01美元		0	
2	CAD {digits}	01加元		0	
3	CHF {digits}	01瑞士法郎		0	
4	EUR {digits}	01欧元		0	
5	FRF {digits}	01法郎		0	
6	GBP {digits}	01英镑		0	
7	HKD {digits}	01港币		0	
8	JPY {digits}	01日元		0	
9	RMB {digits}	01元人民币		0	
10	USD {digits}	01美元		0	
11	VND {digits}	01越南盾		0	
12	£ {digits}	01英镑		0	
13				0	

定义好规则后，退出“系统规则词典”窗口，如下图：当鼠标指向“RMB 300”时就会将它翻译成“300 元人民币”，鼠标指向“EUR 600”时就会翻译成“600 欧元”了。

原文	译文	备注	#	术语
RMB 300	300元人民币			
HKD 400	400港币			
EUR 600	600欧元			
\$500	500美元			
USD 300	300美元			
£ 3.50	3.50英镑			
GBP 230	230英镑			
CHF 750	750瑞士法郎			
FRF 800	800法郎			
CAD 1000	1000加元			
JPY 5000	5000日元			
VND 1000	1000越南盾			

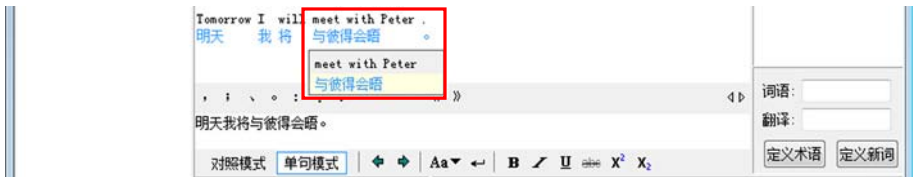
例 4：短语翻译。规则翻译最广泛的应用，是对短语的定义，通过定义规则，使得一些短语可以正确翻译，在翻译“Tomorrow I will meet with Peter. 明天我将与彼得会晤。”这一句时，短语“meet with”默认取词是“遇到”，如下图：但我们希望通过规则定义，日后若遇到“meet with Peter 就能翻译为：“与彼得会晤”，若遇到“meet with Anny 就能翻译为：“与安妮会晤”等等。



为此我们在“系统规则词典”中定义：“meet with \*”的译文为“与#1 会晤”，其中“\*”表示匹配 with 后面的一个单词或短语，“#1”表示翻译第一个“\*”位置的内容。如下图：

序号	原文	译文	类别	优先级	备注
1	meet with *	与#1会晤		0	
2				0	

定义好规则后，我们看到在屏幕取词窗口中，“meet with Peter ”被看成了一个整体短语，其译文为：“与彼得会晤”。

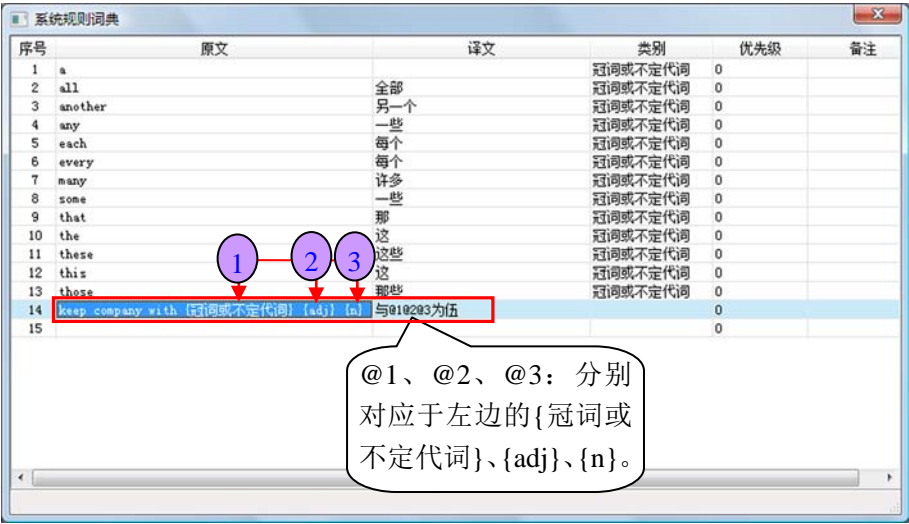


同理，如果我们将“keep company with...”的规则也像刚才那样定义为“keep company with \*”，其译文为：“与#1 为伍”，但发现当翻译“I refused to keep company with those bad boys. 我拒绝与那些坏的男孩为伍。”屏幕取词窗口出现的是“keep company with those 与那些为伍”，如下图：而我们想要的结果是将“those bad boys”看成一个整体，整个短语部分翻译为“与那些坏的男孩为伍”。软件能否做到呢？





由于“\*”表示匹配这个位置的一个单词或一个短语，所以上述中只取了“those”这个单词，而忽略了后面的“bad boys”。我们知道，像一个冠词或不定代词+形容词+名词的这种结构应该看成是一个名词性短语，于是我们将 a、the、this、that、thest、those…… 的类别定义为“冠词或不定冠词”（类别名称可自由定义，{类别名称}表示引用所定义的这个类别）。并在“原文”列中定义“keep company with {冠词或不定代词} {adj} {n}”，“译文”列中输入“与@1@2@3 为伍”，其中@1 表示对应第一个类别“{冠词或不定代词}”的译文，@2 表示对应第二个类别“{adj}”的译文，@3 表示对应第二个类别“{n}”的译文，如下图：



定义好规则后，我们回到刚才翻译的句子，看到此时“keep company with those bad boys”已经是一个合并的短语了，并根据所定义的规则，将它翻译为“与那些坏的男孩为伍”，以后若是遇到“keep company with these good girls”等类似的结构都可以正确翻译。



### 三、优先级

当出现可以匹配多个合适的语法规则时，用户可以设定各个规则的优先级，优先级用 0~20 的数字定义，默认为 0 级，级别越高表示越优先。




### 四、多重嵌套

利用规则中的“类别”，可以进行多重嵌套引用，考虑到匹配速度和实用性，目前软件最多支持 3 重的嵌套。

### 五、语法规则的共享

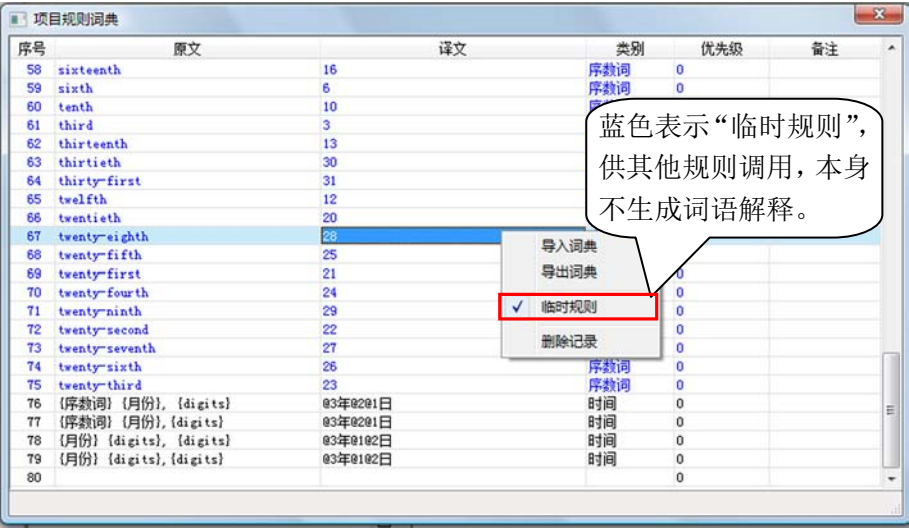
语法规则的整理定义是一项繁琐、耗时的工作，为了减轻工作量，用户之间可以分工合作、互通有无。在“规则词典”窗口中点击【右键菜单→导出词典】功能可以将定义好的词典导出，点击【右键菜单→读入词典】功能可以读入规则，读入的规则会自动排序，为了日后的可读性，在定义规则时可以在“备注列”输入备注，点击“备注列”可以按备注信息排序，利用这个特性，可以将一些不希望被打乱的规则在“备注列”输入特别的标识。

说明：由于在 V1.14 版中增加的语法规则翻译，所以新版本软件保存的文件不能在旧版本软件中打开，如果不希望使用规则词典，点击【项目管理→项目设置】菜单，在“项目设置”窗口中有“不使用规则词典”这一选项，将它打勾即可。

另外，为了便于大家学习，我们将自定义规则的案例及上述所讲的各种规则都一并放在了大家下载的文件夹中，解压后可以看到这两个文件： 自定义规则翻译案例 和  规则样例。其中“ 规则样例”文件夹下面的文件可以导入到“系统规则词典”或“项目规则词典”中。

### 六、临时规则

规则词典增加了“临时规则”的类型，临时规则只用于其他规则中（即是一个中间过程规则，供其他规则调用的），它本身不能直接生成词语解释，临时规则的好处是可以避免各规则间的相互干扰。在规则词典窗口中，选择规则记录，点击【右键菜单→临时规则】功能，软件会以蓝色字体显示，以示区分，如下图。



七、用更简洁的方法定义规则词典

调整了规则词典的格式，使得规则词典的定义更加简洁。增加了“{ :词语 1|词语 2|...|词语 n}” 的表示方法，用于匹配其中之一 的词语（注意：{ }里面要以:冒号开头）。支持了同一个短语的多个解释。译文中可以使用“解释 1|解释 2|...|解释 n” 的方式定义。这样，以前需要用多条规则才能表示的，现在可以用一条规则来定义。下面将举例来说明：

例：我们之前曾经讲述过关于图号规则的定义，为了能将“Figure 1.0、2.0、3.0；Fig-1、-2、-3；Figure 1、Figure 2、Figure3 ”这类表示方法翻译成“图 1.0、图 2.0 或图-1、图-2”等，我们用了 3 个规则来定义它，如下图：



而现在，我们可以用一条规则来定义它，由于“{ :词语 1|词语 2|...|词语 n}”表示匹配其中之一 的词语，于是上述可以简写成：“{ :Fig.|Fig-|Figure} {digits}”，对应的译文写成“图 @2”，如下图：



若同一个词语有多种译法时，译文中的规则可以使用“解释 1|解释 2|...|解释 n”的方式定义，如下图：定义了可能有“图-@2| (图-@2)|<图-@2>|[(图-@2)]图@2”译法。



序号	原文	译文	类别	优先级	备注
1	{:Fig. Fig. Figure} {digits}	图-@2  (图-@2) <图-@2> [(图-@2)]图@2		0	
2				0	

这时，当鼠标指向该词时，会出现屏幕取词窗口，里面列出了刚才定义的多种译法供选择，如下图：（因为涉及“<>[]”等特殊符号，“<图-1>”前面的斜杆“\”软件自动加入，点击输入时“\”不会出现）

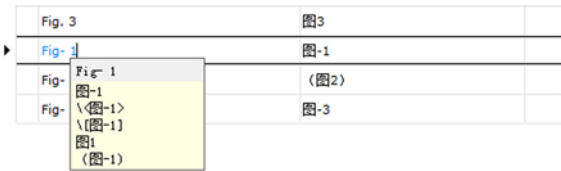


Fig. 3	图3
Fig- 3	图-1
Fig- 1	(图2)
Fig- \<图-1>	图-3
Fig- \[(图-1)]	
Fig- (图-1)	

再如下图中的：“{:is|are|am|be|was|were} born {:in|on|at} 出生于”这样一个规则，就可以将 is born in...、are born on...等等都译成“出生于”了。



序号	原文	译文	类别	优先级	备注
1	{:is are am be was were} born {:in on at}	出生于		0	
2	{Usd USD US\$} {digits}	@2美元		0	
3				0	

另外，还增加了“label:digit\_”和“label:letter\_”类别，“label:digit\_”表示以数字开头的字母组合，例如 1ABC、2XYZ 等。“label:letter\_”表示以字母开头的字母组合，例如 A1B1C1、X2Y3Z8 等等。

在对照模式中，当存在规则词典匹配的词语时，软件使用浅蓝色显示，以示提醒，如下图的“born in June 1, 1999”：



有规则匹配的词语以浅蓝色提示。

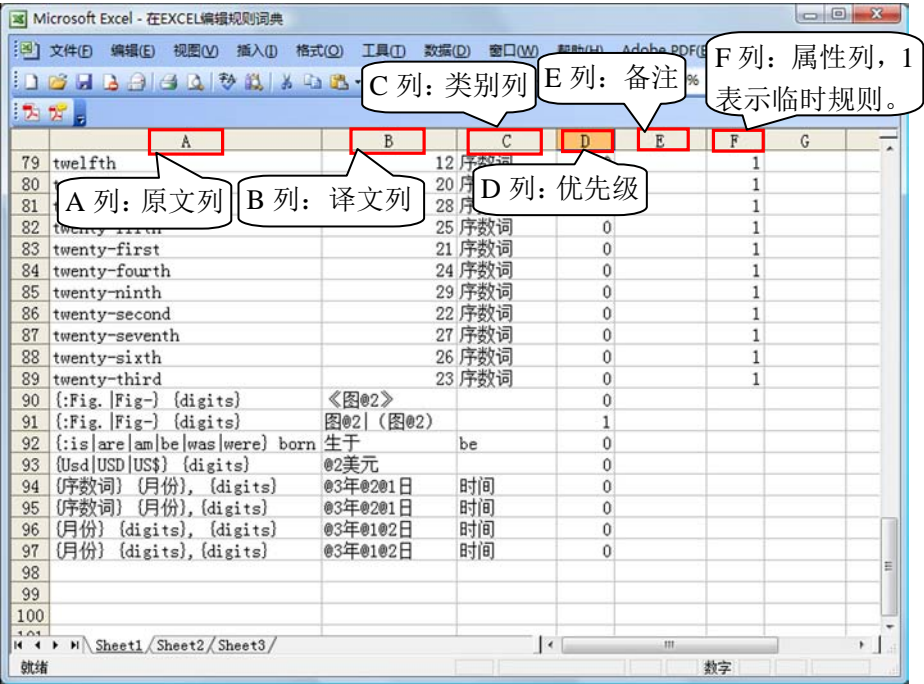
1996年3月2日	
2009	2009年12月8日
26, 1989	1989年1月26日
he is born in June 1, 1999	生于1999年6月1日
born in June 1, 1999	生于1999年6月1日

对照模式 单句模式

词语: 翻译: 定义术语 定义新词

八、在 EXCEL 中编辑语法规则词典

雪人的语法“规则词典”功能强大，定义起来简洁、灵活，深受不少用户的喜爱！但若进行大规模的语法规则定义，在软件本身的“词典规则”窗口，由于屏幕空间等因素的限制，不是十分的方便。为此，雪人在 V1.19 版的规则词典中增加了导入、导出 EXCEL 文件格式的支持，方便用户进行大量的语法规则定义，定义好后，在“词典规则”窗口中点击【右键菜单→导入词典】功能，将它导入到雪人的“规则词典”中。在 EXCEL 中编辑时，各列的格式应该与规则词典窗口的内容相对应，即 A 列定义为：原文列；B 列定义为：译文列；C 列定义为：类别列；D 列定义为：优先级；E 列定义为：备注；F 列定义为：属性列“1”表示临时规则。如下图：








另外，雪人在上一版本规则词典中，就增加了“{:词语 1|词语 2|...|词语 n}”用于匹配其中之一词语。“解释 1|解释 2|...|解释 n”用于对同一个短语的译文定义多个解释。现在，对“类别”也支持类似的表达方法，“{类别 1|类别 2|...|类别 n}”用于匹配其中之一类别，这样就使得定义起规则来更加的简洁和方便！

提示：若将规则词典导出到 EXCEL 中时，注意所定义的规则中，避免以“=”字符开头，因为“=”在 EXCEL 中被赋予了编辑公式的应用，若规则以“=”字符开头，则导出时会报出错。

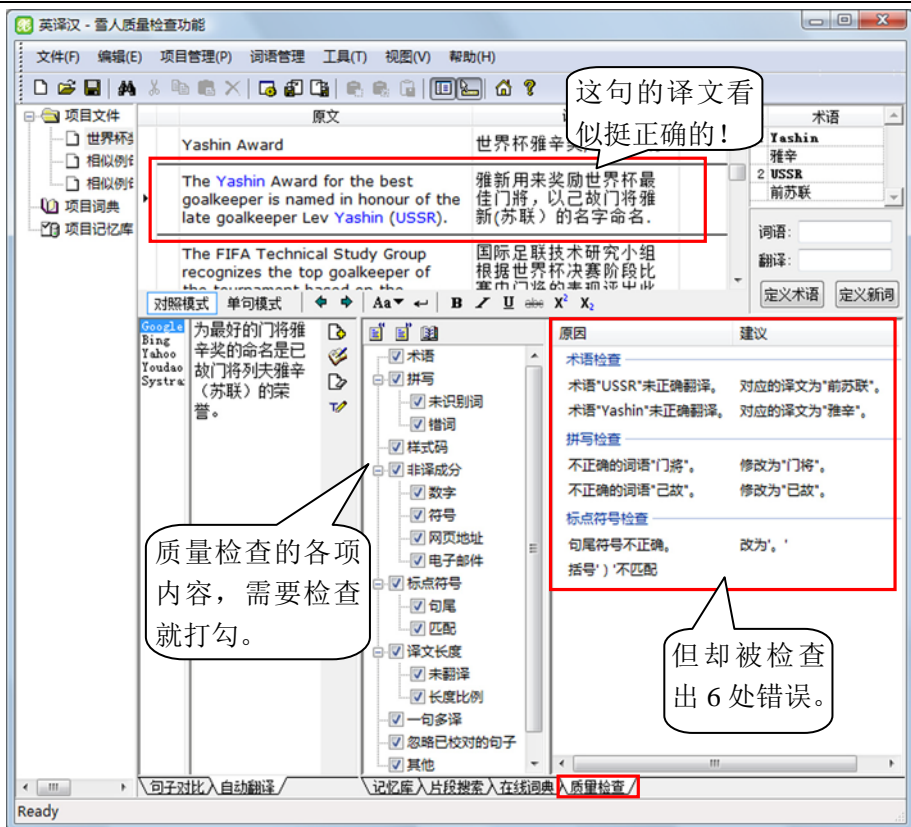
## 第十四章 质量检查

雪人提供了非常强大的“质量检查”功能，质量检查的项目具体包括：1. 术语检查；2. 未识别词和措辞的拼写检查；3. 样式码检查；4. 数字、符合、网页地址、电子邮件等非译成分的检查；5. 标点符号的检查；6. 未翻译的句子、句子长度的译文长度比对检查；7. 一句多译的检查等等。雪人的质量检查功能可以帮助译员从一些细微但极其重要、且往往令人疲于应付的翻译的“小事”中解脱出来，使你的译稿质量更有保障！

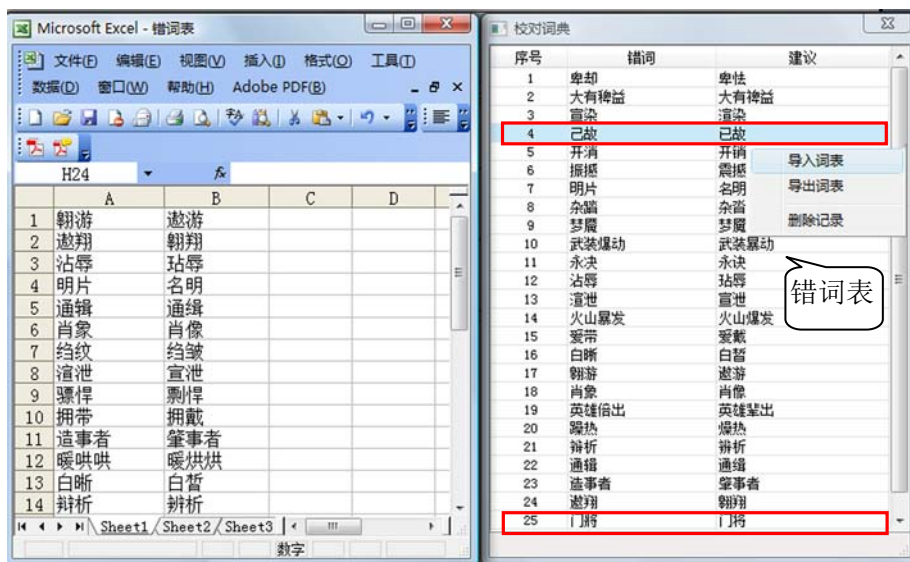
点击屏幕下方的“质量检查”页签，可以看到该页面窗口的内容，如下图所示，其中有 3 个按钮，分别是： “向前检查有问题的句子”， “向后检查有问题的句子”， “校对词典”。建议用户在完成整篇的译文翻译后，再切换到这个“质量检查”页签，通过点击 “向后检查有问题的句子”按钮来进行质量检查。雪人的质量检查是逐句检查有问题句子，光标定位在有问题的句子上，供用户马上对该句的错误进行修改，修改完毕后，再点击 按钮，直接跳到下一个有问题的句子。

下面我们看看下图中的这个例句：“The Yashin Award for the best goalkeeper is named in honour of the late goalkeeper Lev Yashin (USSR). 雅新用来奖励世界杯最佳门将，以已故门将雅新(苏联)的名字命名。”乍看这句译文好像挺好的，但是当我们切换到质量检查窗口时，雪人却检查出该句存在 **6 个错误**！哪 6 个？分别是 2 处术语错误，2 处错别字，2 处标点符号错误。原来，我们将“Yashin 雅辛”和“USSR 前苏联”这两个词定义为了术语，但在实际翻译中，我们确将它们翻译成了“雅新”和“苏联”，所以软件的“术语检查功能”提示术语没有正确翻译，并建议更改为“雅辛”和“前苏联”。2 处错别字提示是“门将”改成“门将”，“已故”改成“已故”。2 处标点符号提示是：句末应该用中文的“。”，而我们却用了英文的“.”，还有是“（苏联）”这个对括号不匹配的问题。

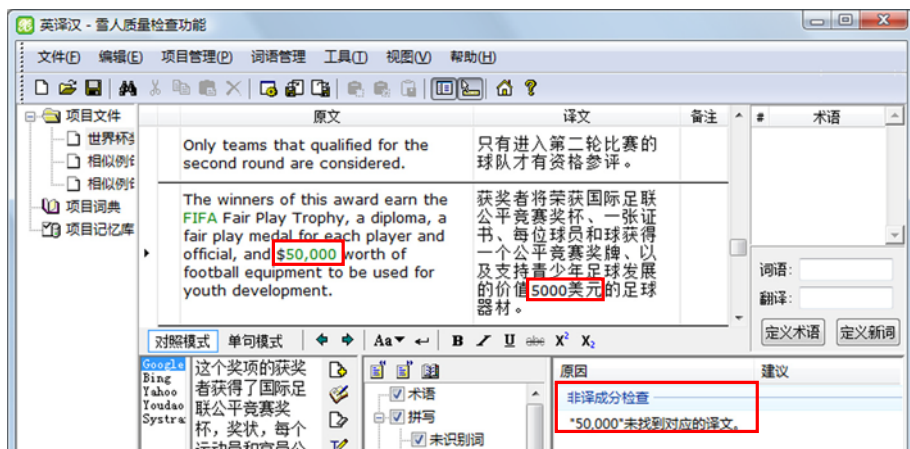




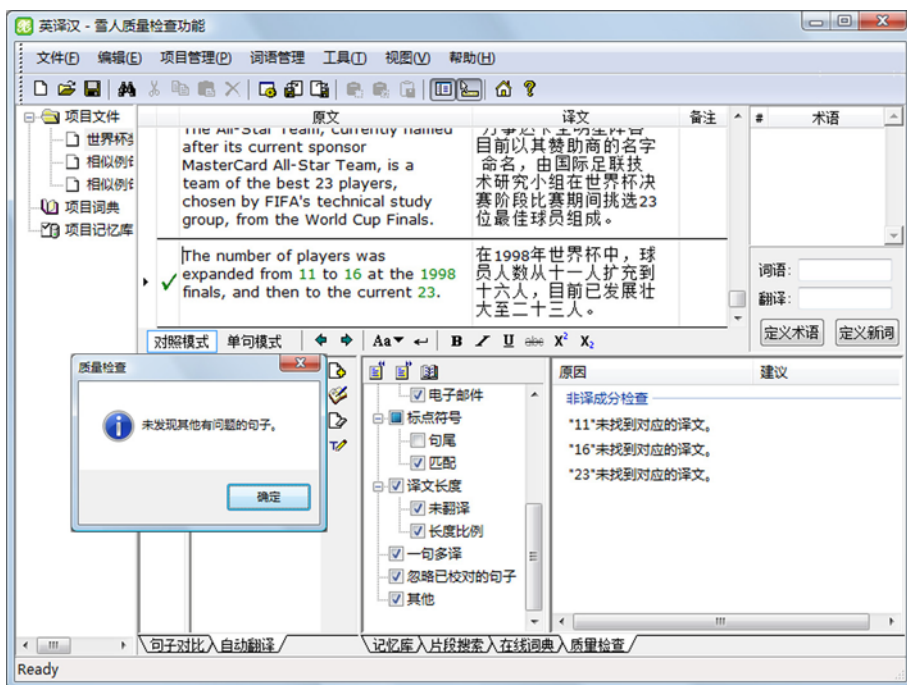
雪人对错别字的检查，得益于校对词典中的错词表，点击 “校对词典”按钮可以看到，这个错词表用户可以自由添加、删除，英文的会被自动保存在安装目录下的“english.dct”文件中，中文的则保存为“chinese.dct”。因为通过词表进行中文校对的效率不高，在默认情况下，默认情况下标准版不会自动安装这个词表文件。标准版的用户可以登录[http://gcys.cn/fy\\_xiazai.html](http://gcys.cn/fy_xiazai.html) 下载中文校对词表，下载后将它解压到安装目录\SCAT下面，并重新启动软件即可。免费版的用户可以按需录入或者导入EXCEL文件建立“校对词典”。建立的方法是：1. 直接在“校对词典”窗口中添加；2. 或在EXCEL中分正确的、错误的两列，编辑好后导入。如下图：



当我们修改完这上一句的所有错误后，点击 按钮会直接跳到下一个有问题的句子（没有问题的句子会直接被跳过）。如下图：软件发现数字“50000”没有对应的译文，仔细一看，原来是在译文中将“50000”译成了“5000”。

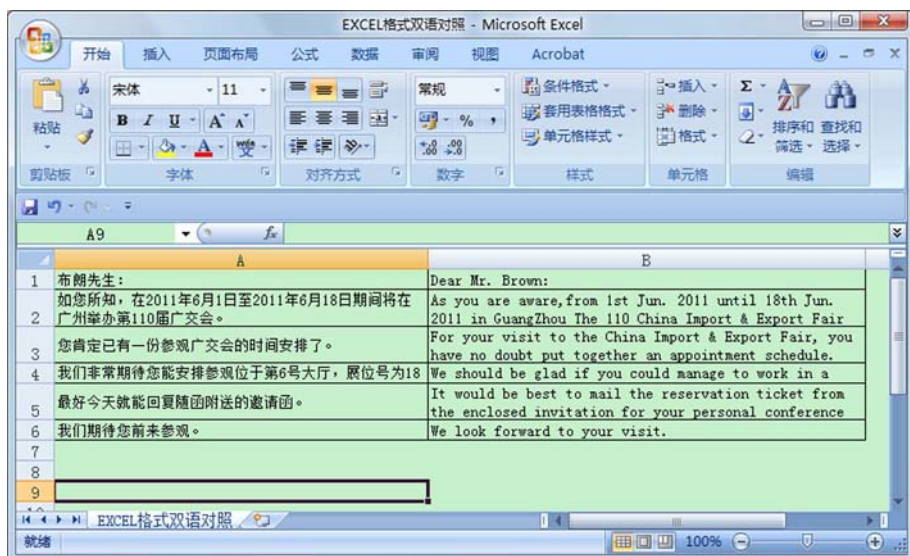
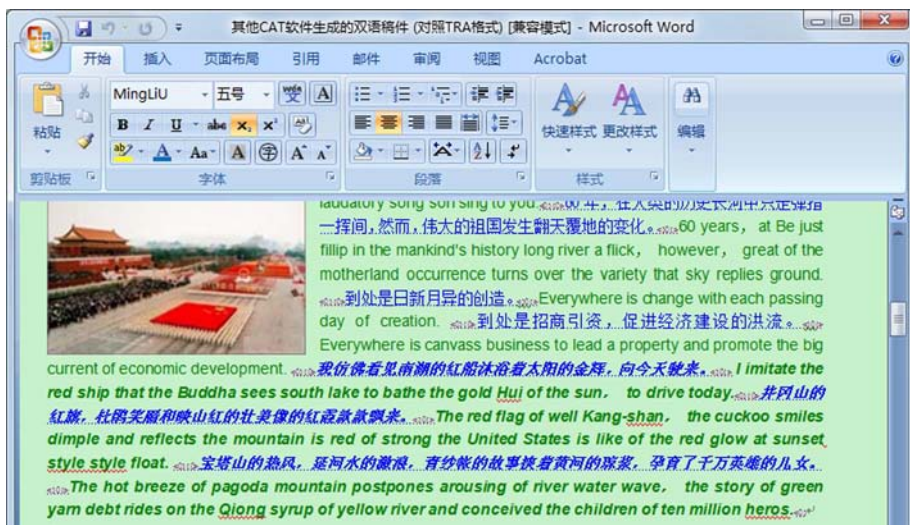


修改好后，继续按 按钮跳到下一个有问题的句子，如下图：这时提示“11、16、23”数字没有对应的译文，其实我们是将它们翻译成了汉字“十一、十六、二十三”，经校对后认为这不是错误，此时需要按【F9】键，这时会出现“√”，表示该句是正确的，下次不需要再跳到该句检查。当所有错误检查完毕时，再按 按钮时，会跳出“为发现其他有问题的句子”提示窗口，此时表示检查完毕，可以放心地交稿了！



## 第十五章 利用雪人 CAT 进行审校

如果其他译员不是使用雪人 CAT 进行翻译的,或者是直接翻译的,能否利用雪人 CAT 对他们的稿件进行审校呢?回答是肯定的。雪人的“导入双语文件”功能,可以很方便地校对其他 CAT 生成的 Unclean 格式的双语文件或 EXCEL 的双语文件,例如下面两图:

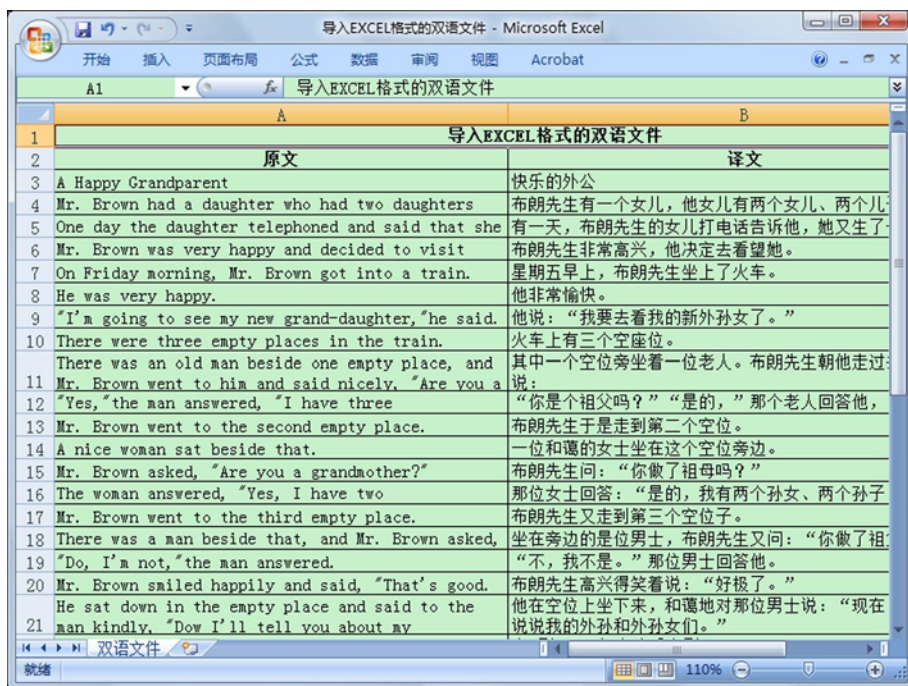


右键点击屏幕左侧的“项目文件”，然后点击【右键菜单→导入双语文件】功能，将上述文件导入，导入后就可以在雪人中进行审稿，软件会自动将这些双语句对保存在“项目记忆库”中，由于导入的双语文件不包含段落信息、字体样式、图片等等信息，若审校完毕后就直接导出译文，则仍需重新排版。利用雪人的“预翻译”功能可以解决这个问题，方法是：将需要翻译的原始文档导入，然后利用刚才形成的记忆库进行快速的“预翻译”（点击【工具菜单→预翻译】功能）自动填写译文（部分断句不一致的句子可以进行手工断句）。然后重新导出译文，就有了与原文风格一致的译文了。



如果收到的审校稿件是 TXT 文本文件、WORD 格式的文件，又该如何在雪人中审校呢？方法还是有的，首先利用雪人的双语对齐功能，将 TXT 文本文件或 WORD 文件进行句与句的对齐，对齐好后，点击【右键菜单→导出双语文件】功能，将它导出成 HTML 格式的文件，这个 HTML 文件也可以用 EXCEL 软件打开它，如下图：





接着回到雪人 CAT 中，在屏幕左侧的“项目文件”中点击【右键菜单→导入双语文件】功能将这个 HTML 文件导入即可进行审校了，结果如下图：





## 第十六章 网络协同翻译平台

### 第一节 客户端与服务器连接

雪人 CAT 协同翻译平台”简化了服务器架设及对服务器系统的要求，无论是拥有专门的技术人员、高性能的服务器的大型企业，还是小型的翻译公司和翻译团队都可以轻松地架设雪人 CAT 协同翻译平台，使翻译企业、翻译团队**拥有自主的协同翻译平台**，切实保障翻译资料的安全、保密性。“雪人 CAT 协同翻译平台”由服务器端软件和客户端软件构成。不仅可以在 Windows Server 2003/2008 服务器版本的操作系统下使用，也可以在 Windows XP 及以下的操作系统下使用。在雪人官方网站，有免费版的服务器端软件下载。

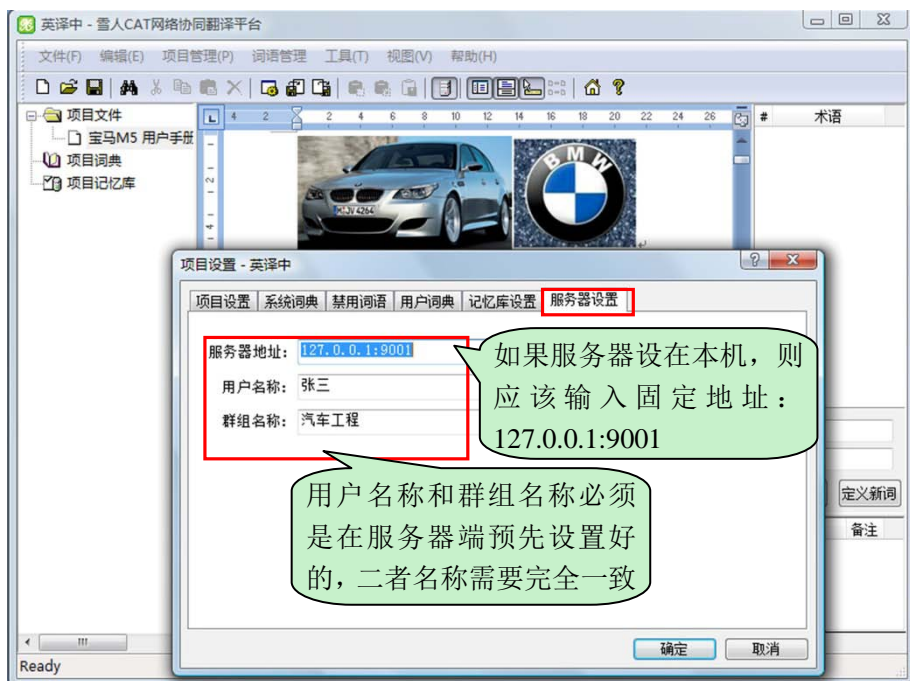
雪人 CAT 软件既可以单独使用，也可以作为协同平台的客户端软件使用。雪人 CAT 软件在 **V1.27 版中增加了连接服务器的相关功能**，用户只要进行简单的设置，就可实现多人协同翻译了。


下面以“汽车工程”翻译项目为例，介绍如何进行多人协同工作。

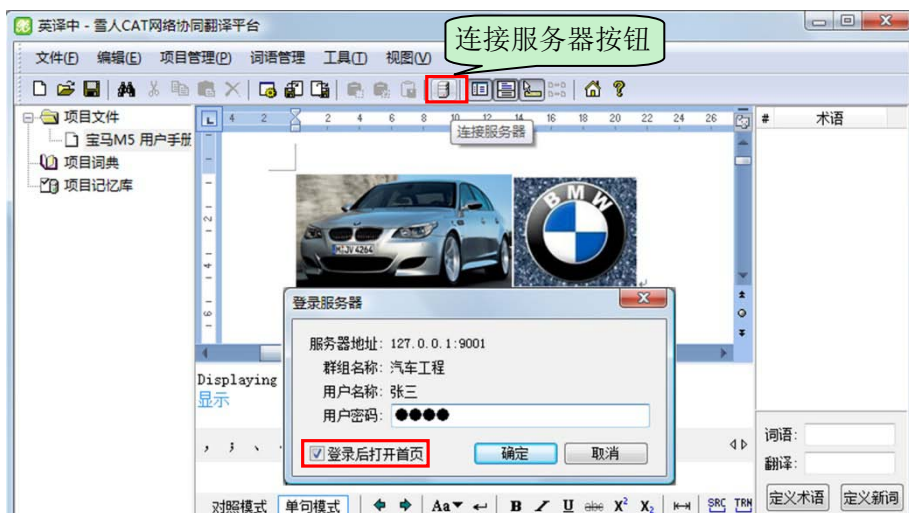
汽车翻译行业具有大量的重复性的产品名称和专业术语。主要翻译文档有：产品技术标准、工艺标准、材料标准、用户手册、维修手册、零部件目录、培训教材等等众多资料。若单兵作战难以在短时间内完成如此大量的翻译工作，因此多人分工协作就成为了必然。**如何保证大量重复性的专业术语和产品名称在同一项目或多个项目中的统一，成为了影响项目翻译质量的关键因素。**雪人 CAT 的网络协同翻译平台可以很好地解决这个问题。

在服务器中，我们创建了“**汽车工程**”项目组，该项目组由 5 人组成，并设置了：张三为项目组长，李四为审校，王五、杨辉、王小明为译员。下面我们以“项目组长 张三”的身份讲述如何将客户端软件与服务器相连接，实现协同翻译。

启动雪人客户端软件，新建翻译项目，导入需要翻译的文档，然后点击【项目管理➡项目设置】菜单，V1.27 版后，在“项目设置”窗口中增加了**服务器设置**页签（如下图），输入拟连接的服务器的 IP 地址及端口，例如：192.168.1.88，**如果服务器是设在本机的，则输入固定的 IP 地址 127.0.0.1**（如果服务器的端口不是 9001，需要在地址后面添加端口号，如 127.0.0.1:1234）；输入用户名称及需要参加的群组名称，**用户名称和群组名称必须是在服务器端预先设置好的，二者名称需要完全一致。**



点击“确定”按钮后，点击工具栏的“ **链接服务器**”按钮，在弹出的“登录服务器”窗口中输入你的登录密码，当前的翻译项目就与服务器建立好链接了。**连接服务器后，翻译句子时会同步到服务器中，同项目组的其他成员如果遇到与之相近的句子时，可以同步显示在他的客户端软件的记忆库提示窗口。**审校人员也立即看到你翻译的句子、定义的术语，在确认无误后可以审核入库，如果有问题可以已邮件方式通知用户更正。术语经审核后，同项目组的其他成员可以即时共享。



如果勾选“**登录后打开首页**”，则按“确定”按钮后，会打开服务器端的用户首页。如果不需要查看服务器中的信息，可以不勾选这个选项。登录服务器**首页**的界面如下图：如果有新邮件会在首页中提醒你，点击可以查看邮件的内容。项目组长的新布置的任务，平台成员最新发表的文章、帖子等等都会第一时间通知你，点击可以查看内容。



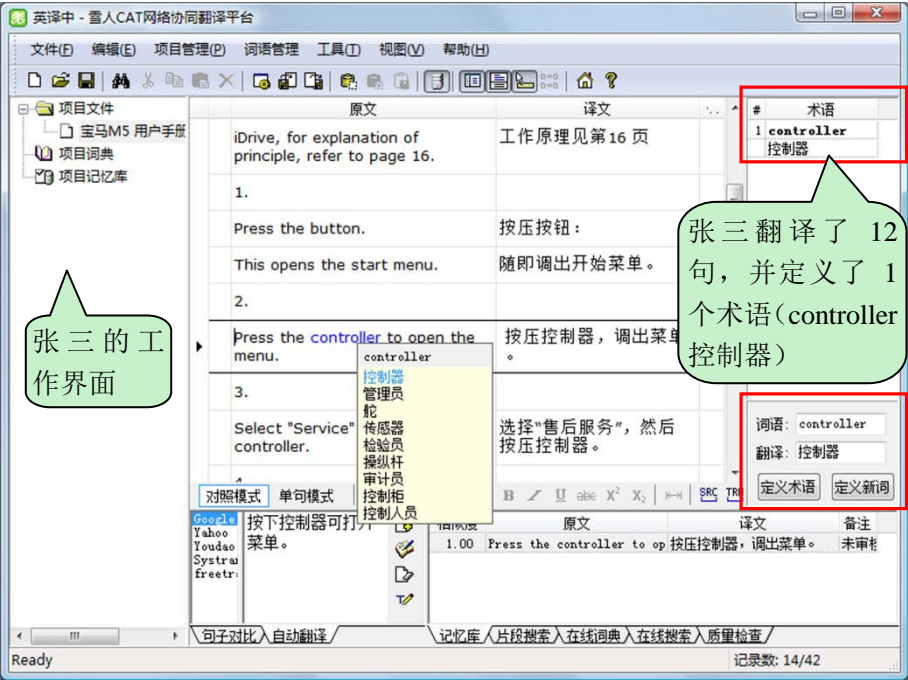
点击导航栏的“项目”按钮，然后点击左侧“【项目管理→汽车工程】”项目组（你参加了多少个项目组，在屏幕左侧就会显示所参加的项目组名称），如下图。



上图中我们看到新创建术语库和记忆库的记录全部显示为 0，那是因为该项目组的成员还没有开始实际的翻译工作。一旦该项目的成员翻译一句或增加一个术语，就会立即显示在此。**管理员和审校人员看到的是项目的合计记录数，译员看到的是自己建立的记录数。**

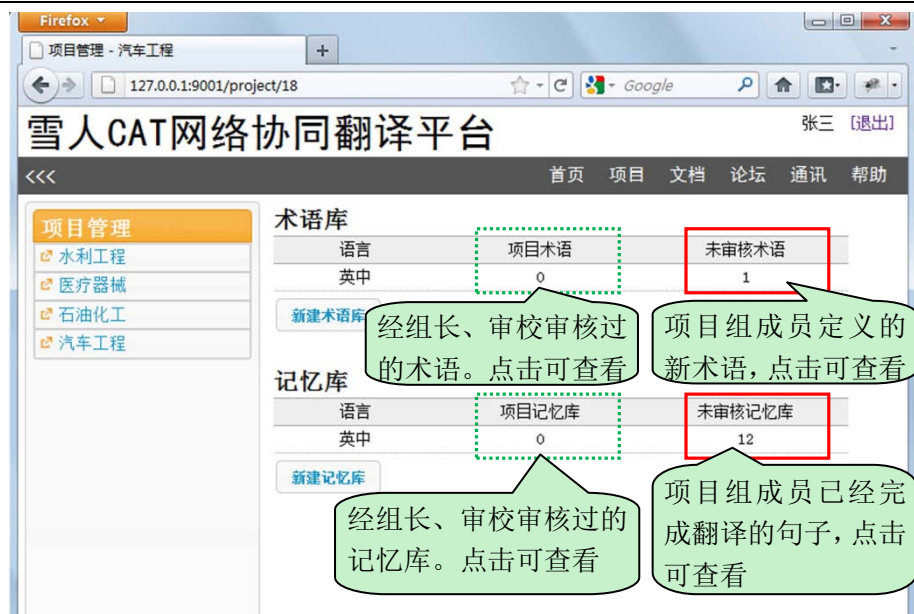
## 第二节 实时共享记忆库、术语库

下面回到客户端软件，进行实际的翻译工作。例如下图中，张三翻译了《宝马 M5 用户手册\_第1部分》的前12句。由于“controller”在词典中有众多译法，张三将“controller”翻译成“控制器”，并将它定义为术语。



此时我们再切换到服务器的“项目”页面中，我们看到张三刚才翻译的 12 个句子和 1 个术语已经自动上传到服务器中了。界面如下图：

如果用户的身份是普通译员，则“未审核术语”显示的数量是自己新定义的术语；如果是项目组长、审校则“未审核术语”显示的数量是本项目组所有成员新定义的术语数量，**点击数字可以查看具体的内容**，“未审核记忆库”数量也类同。“项目术语”和“项目记忆库”用于存放审核过的术语和句子。经过审核的术语，客户端软件可以同步使用。



点击“未审核记忆库”的具体数量，点击“数量 12”，可以查看新翻译的句子的内容（如下图）：审校可以对这些句子进行审核。对翻译的正确的句子，就在选择框中打勾，然后点击下方的“核准记录”按钮审核入库，**经审核的句子会自动转入“项目记忆库”中**，并在“项目记忆库”列中显示具体的数量，点击“具体的数量”可以查看。**经审核和未审核的记忆库中的句子全项目组的所有成员都可以使用，不同之处是：未审核的句子会在客户端的“备注”列标注“未审核”字样，以示提醒。当有多个相似句子时，审核了的句子会优先排在前面。在使用预翻译时，软件只选择审核通过的句子。**



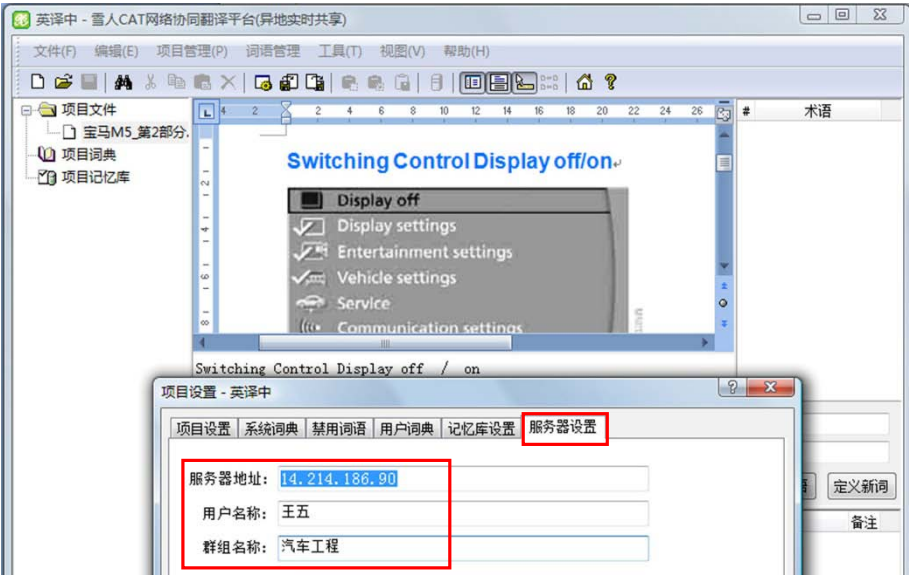
如果觉得屏幕空间不够，点击导航栏左侧的“<<<”按钮，可以将左侧菜单隐藏，此时按钮会变成“>>>”，点击它又可以再次显示左侧菜单，如下图。



点击“未审核”术语的数值，可以查看具体的术语内容，例如下图中，张三在客户端中定义的术语“controller 控制器”就自动上传到服务器了，由于该术语还没经审核，所以在项目组的其他成员的客户端软件中还不会出现。术语只有在审核通过后，项目组的其他成员才能使用。

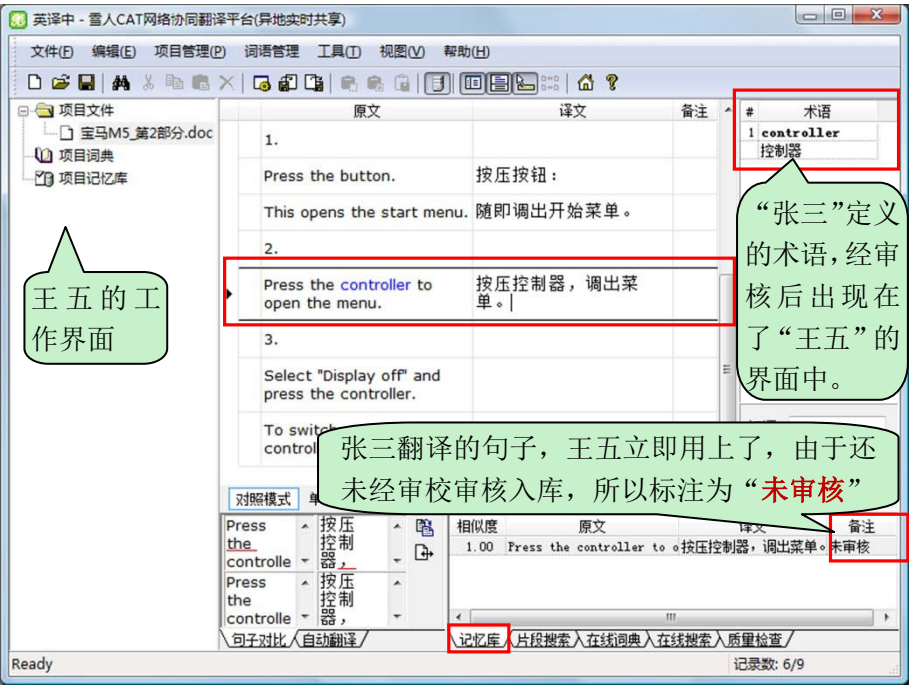


如果已经将张三定义的术语审核入库，并将他翻译的部分句子也审核入库。接下来我们看看张三所完成的工作对同项目组的异地成员带来的帮助。北京的王五负责翻译《宝马M5 用户手册\_第2部分》。因为服务器是设在张三的机器上，所以张三访问服务器时输入127.0.01:9001 这个本机固定访问地址。而王五则不能使用这个地址，张三需要将自己机器的IP地址告诉王五，王五才能登录。经查询，张三的服务器机器地址是：14.214.186.90，则王五在客户端软件的“服务器设置”中输入如下图：



点击“确定”按钮后，点击工具栏的“[链接服务器](#)”按钮（再次单击该按钮，可以断开与服务器的链接。），在弹出的“登录服务器”窗口中输入登录密码，则“王五”当

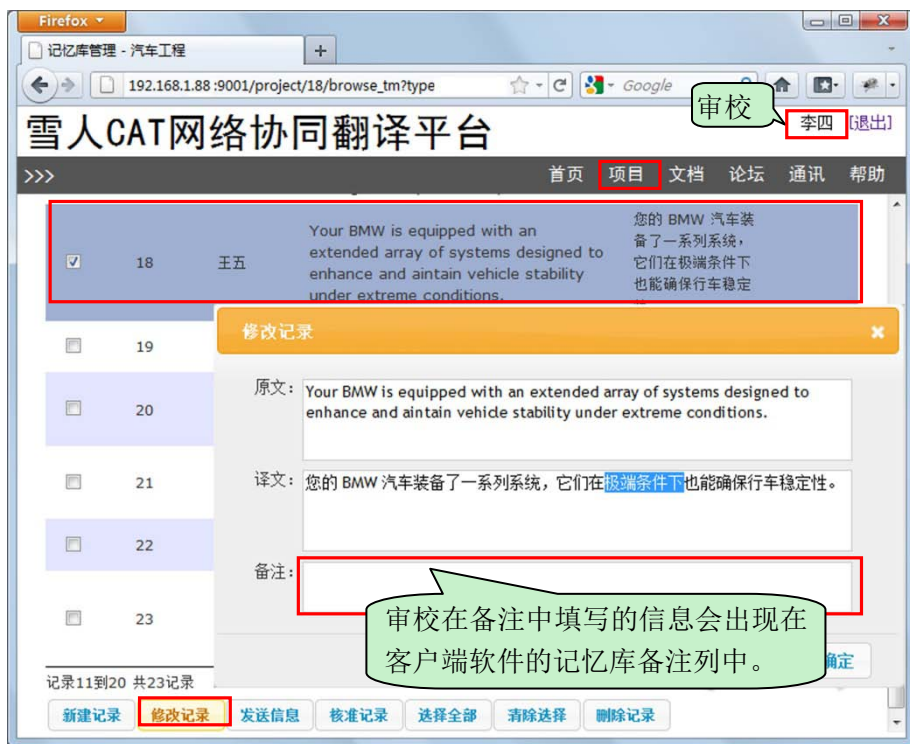
前的翻译项目就与服务器建立好链接了。当王五在翻译“Press the button. This opens the start menu. Press the **controller** to open the menu.”这几句时，由于张三已经翻译过，光标去到该句时，根本就不用翻译而直接上屏了。遇到“**controller**”这个词时，在“王五”的工作界面的屏幕右侧自动出现了该术语。如下图：



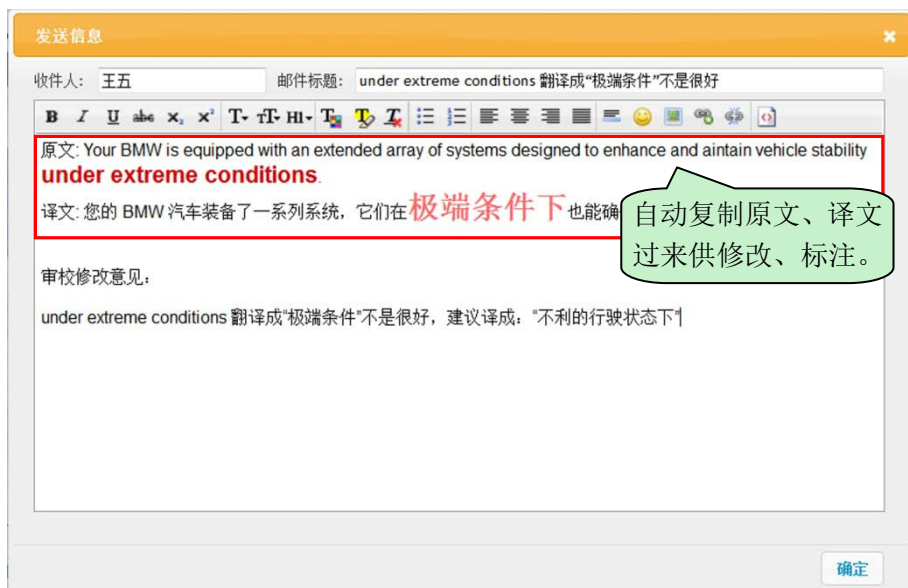
### 第三节 审校与翻译同步

利用雪人 CAT 协同翻译平台，通过互联网技术可以连接异地的译员、审校，实现在译员翻译的同时审校同步对翻译稿件进行审核，将传统的译员审校的流水线工作模式转变为并行工作模式，大大节约了整体翻译工作时间。审校可以及时发现译员存在的问题，及时纠正，达到“一人问题、及时解决、全体收益”的目标，避免等到译员初译工作完成后才发现而造成的诸多麻烦。


现在，审校“李四”直接在浏览器的地址栏中输入服务器的地址登录服务器（注：为了安全起见，利用网页登录服务器的，若不进行操作，超过 15 分钟后会自动退出。利用客户端登录的不受此限制），在“**项目**”页面中，对各地译员翻译的译文进行审核。在审核“王五”翻译的内容时，发现他将“under extreme conditions ”翻译成“极端条件”，感觉这样翻译不是很好。翻译成：“不利的行驶状态下”会更加贴切些。此时，审校可以选择该句后，点击下方的“**修改记录**”功能直接修改，修改好后再入库。在“备注”中填写的信息会出现在译员们的客户端软件的记忆库备注列中，审校可以根据需要确定是否需要为该句填写备注。

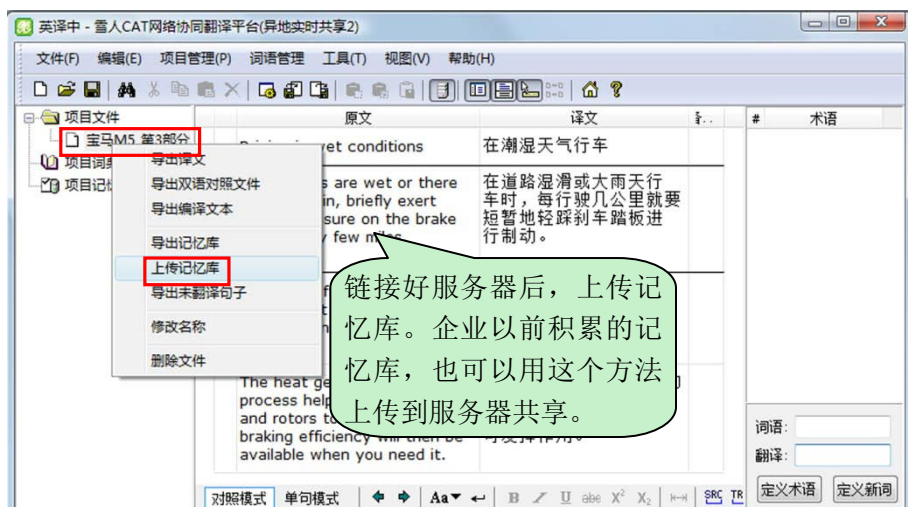


如果觉得必要，审校还可以将该句存在的问题发信息通知译员“王五”。点击屏幕下方的“**发送信息**”，软件会自动将该句的原文、译文复制在“发送信息”的编辑窗口中，并自动将“王五”设为收件人。审校可以在编辑窗口中标注错误的地方，填写修改的意见等等。如下图：

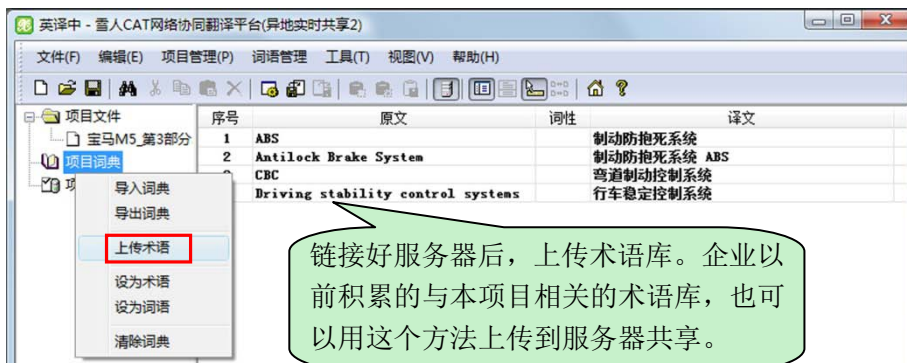


这样,“王五”在登录服务器后,就可以收到审校“李四”发来的邮件了。

译员“王小明”负责《宝马 M5 用户手册\_第 3 部分》的翻译,在翻译时由于某种原因并没有连接服务器。他该如何上传自己的翻译内容给审校审核呢?首先,向前面的介绍一样,在客户端软件的“服务器设置”中设置好,然后点工具栏的“链接服务器”按钮,链接好服务器。再点击屏幕左侧的“项目文件”下面具体需要上传的文件,选择【右键菜单→上传记忆库】功能,将当前文件的记忆库上传到服务器中,如下图。

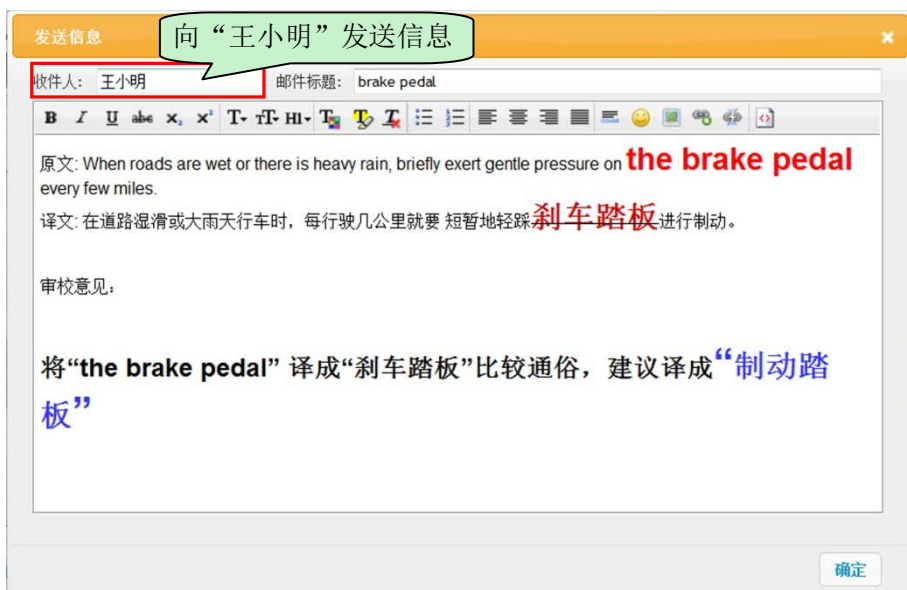


术语的上传方法也类似，右键点击屏幕左侧的“**项目词典**”，选择【**右键菜单→上传术语库**】功能，将翻译过程中定义的术语上传到服务器给审校审核。**企业以前积累的记忆库、术语库也可以用这个方法上传到服务器共享。**

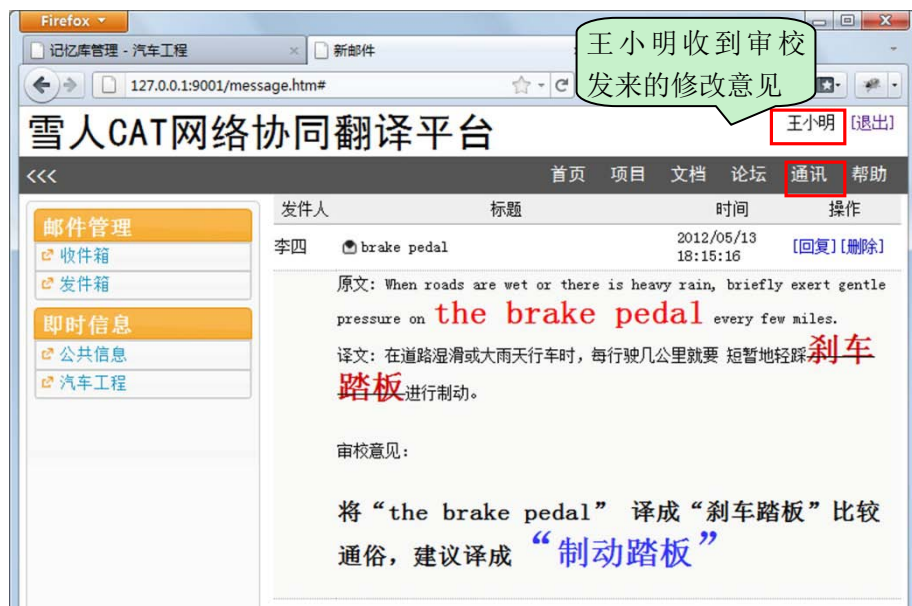


上传后，审校“李四”就能对他翻译的内容进行审核了。在审核时，发现“王小明”将“the brake pedal”译成“刹车踏板”觉得不够好，此时审校可以自己直接修改，但也可以不修改，而是点击屏幕下方的“发送信息”按钮，将修改意见发回给“王小明”，由他自行修改。如下图：

“王小明”在登录服务器后，会收到来自审校“李四”信。如下图：







“王小明”根据信中的内容，在客户端软件中修改该句，修改好后，服务器端的该句会同步更新，如果用户在客户端软件中对译文进行了多次修改，服务器中的该句会以最后修改的版本为准。

更多关于网络协同翻译平台的介绍，请参看服务器端软件中附的操作手册，这里仅摘录了客户端与服务器连接的相关内容。

# 第十七章 版本区别及配置要求

为了满足不同译员对软件的不同需求，雪人翻译同时推出两个版本：免费版和标准版；虽然对免费版的功能有所限制，但**免费版亦能满足一般的翻译需求，可以应用于实际的翻译工作中，而且免费版一样会不断增加新功能，不断升级。**标准版是需要付费使用的版本，标准版的功能相当强大，可以更好地提高您的翻译工作效率，如果您是一名专职翻译工作者，那么它对您来说绝对物超所值！

## 一、免费版和标准版的功能区别

	免费版	标准版
对原文件格式的支持不同	只支持翻译 TXT 文件	支持翻译 WORD、EXCEL、PPT 等多种格式的文件
系统自带词典的不同	系统自带 30 万词的基本英汉、汉英词典	系统自带 60 多个专业、1000 多万的英汉、汉英词典
对导出记忆库格式的支持不同	只支持导出 STM 雪人格式的记忆库	支持导出多种格式的记忆库，如 STM、TMX、EXCEL、HTML 等等
对导出词库格式的支持不同	只支持导出 DIC 雪人格式的词库	支持导出多种格式的词库
对词形变化的处理不同	不持 词 形 变 化 ， 如 ： would/will, apple/apples 会认为是两个不同的单词	自动处理词形变化，提高匹配率，使得记忆库的作用得以更好的发挥
机器自动翻译	不提供机器自动翻译的译文作参考	对每一句都提供机器自动翻译的译文作参考，译员可以直接选用或修改后在使用
质量检查	不带拼写检查的字典	带拼写检查的词典
待增强的功能不同	对标准版中新增的功能，部分酌情考虑对免费版开放	会不断增强软件的功能
服务支持	有限服务支持	全面的服务支持

## 二、32 位和 64 位的区别

术语 32 位和 64 位指的是计算机处理器（也称为 CPU）处理信息的方式。64 位 CPU 可以有更大的内存管理能力，因此电脑可以使用更多的内存，从而大大提高内存密集型应用的效率。那么 64 位电脑的具体处理速度究竟能有多快呢？专家对此解答说：“当计算机面临大量的数据流时，32 位的寄存器和指令集不能及时进行相应的处理运算。32 位处理器一次只能处理 32 位，也就是 4 个字节的数据；而 64 位处理器一次就能处理 64 位，即 8 个字节的数据。形象地说，32 位好像是一个狭窄的单车道，当车流过多的时候，就无法承载这些负荷，而 64 位好比一个宽阔的高速公路，所以在多任务，多程序处理的情况下，应用 64 位计算平台才能随心所欲的加速、把电脑的应用性能发挥到最好。32 位 Windows 能够支持最多 4 GB 系统内存，每个处理器可以使用最多 2GB 专用内存；而在 64 位 Windows 下，最多可以使用 128 GB 的物理内存和 16 TB 的虚拟内存，相对 32 位 Windows，我们明显的一个感受就是内存数量的支持得到了改善。

由于不同译员拥有的翻译记忆库大小不同，我们将标准版划分为：32 位版本和 64 位版本。用户应该根据自己记忆库大小及机器配置来考虑购买 32 位版本或 64 位版本。雪人 32 位版本可运行在 32 位或 64 位操作系统上，64 位版本只能运行在 64 位操作系统上。如果您需要使用巨型的翻译记忆库（大于 100 万句对以上），建议使用 64 位版本，64 位可以支持上千万的句库。

查看机器是 32 位还是 64 位的方法：法一、在我的电脑窗口单击右键，选择“属性”功能就可以看到了。法二、右击【我的电脑】，选择【右击菜单→属性】，在【系统信息】提示里应该有操作系统信息，会显示你的操作系统位数。

## 三、雪人 CAT 对机器的配置要求

雪人对机器的最低配置要求：1G 以上的内存，32 位操作系统建议使用 2G 以上的内存。

32 位标准版：可运行在 32 位或 64 位操作系统上，支持 XP、VISTA、和最新的 WIN 7。



64 位标准版：只能运行在 64 位操作系统上，支持 VISTA 64 位、WIND7 的 64 位 操作系统。

## 第十八章 常见问题解答


### 1. 我该选购 32 位版本还是 64 位？

答：根据你的记忆库大小及机器配置来考虑。雪人 32 位版本可运行在 32 位或 64 位操作系统上，64 位版本只能运行在 64 位操作系统上。普通用户选用 32 位版本就可以了，雪人的 64 位版本可以**支持上千万的句库**。查看机器是 32 位还是 64 位的方法：法一、在我的电脑窗口单击右键，选择“属性”功能就可以看到了。法二、右击【我的电脑】，选择【**右击菜单→属性**】，在【系统信息】提示里应该有操作系统信息，会显示你的操作系统位数。





### 2. 如何开始翻译工作？



答：标准版需要插好软件加密狗，输入安装序列号，安装好软件后启动雪人软件（免费版不需安装，解压后直接双击  translation.exe 启动软件）；然后点击新建按钮，建立一个新的翻译项目，右击屏幕左侧的  **项目文件**，利用【**右击菜单→导入文件**】功能导入待翻译的文件（免费版只能导入 .txt 文件，标准版是支持多种格式）。

### 3. 我用双语对齐做的记忆库不能导入怎么办？

答：标准版可以导入，免费版不能导入 STM 格式的记忆库，但这不影响免费版用户的使用。**并不是只有将记忆库导入才能用的，通过“添加引用”更节省机器资源！**双语对齐后导出 STM 格式的记忆库文件，日后在任何翻译项目中都可以使用它。“添加引用”引用方法：法一、在新建项目时会自动弹出【项目设置】窗口，点击  页签，点击“**添加**”按钮，选择导出的 STM 格式的记忆库文件即可添加；法二、点击【**项目管理→项目设置**】菜单。添加引用后，你可以利用“片段搜索”功能，随便划选几个词，就可以看到重记忆库找搜索出来的结果了。一般情况下不建议用户将记忆库导入进来，而是建议使用“添加”引用的方式，导入只是为了合并多个记忆库及编辑修改之用。

### 4. 词典、记忆库的添加引用与导入有什么不同？

答：尽量通过“添加”引用外部的词典和外部记忆库的方式来为当前翻译项目指定使用的词典和记忆库，而不是将它导入到  **项目词典** 或  **项目记忆库**。因为“添加引用”的方式可以更好地减少项目文件的大小和对系统资源的占用，对配置不高的机器尤其重要。“添加”引用方法：法一、在新建项目时会自动弹出【项目设置】窗口供用户添加；法二、点击【**项目管理→项目设置**】菜单。屏幕左侧的  **项目词典** 或  **项目记忆库** 用于保存在翻译过程产生的词汇或用户导入的词典、术语或记忆库，一开始为空，用户可以点击【**右击菜单→导入**】功能进行导入或在翻译过程中积累；例如对于客户给定的术语，

可以导入到项目词典后，利用【**右击菜单→设为术语**】功能将它们设定为术语。翻译完成后，应该对  **项目词典** 和  **项目记忆库** 整理后导出，日后通过“添加”引用的方式为其他翻译项目使用。

## 5. 已经添加了记忆库，怎么在“项目记忆库”还是空白？

答：正常。因为“项目记忆库”是用于自动记录在翻译过程产生的句子和用户导入的句库。为了使软件运行的更加快速，“添加引用”的记忆库不会出现在“项目记忆库”中，但这丝毫不影响对“添加引用”记忆库的使用。若翻译时遇到与“添加引用”记忆库中的相似句子，一样会自动提示出来。你可以利用“片段搜索”功能来测试一下，随便划选几个词，就可以看到从“添加引用”的记忆库中搜索出来的结果了。

## 6. 雪人可以导入哪些格式的词典和记忆库？

答：导入词典的格式：Dic(雪人格式)、XLS、XML、TXT、CSV 等格式。导入记忆库的格式：STM(雪人格式)、TMX、XLS、TXT、CSV、HTML 等格式；文本格式的导入非常灵活，既可以自动识别分隔符，也可以手工指定分隔符。在导入文本格式时，软件会自动识别分隔符，并将识别到的分隔符提示给用户，用户可以修改；如果原文中存在分隔符不统一的情况，则软件不能正确识别，这时需要用户在提示窗口中手工输入分隔符。

## 7. 如何制作 EXCEL 格式的句库或词典导入到雪人中？

答：在 EXCEL 中分两列：一列中文、一列英文就可以导入到雪人中了。**记忆库是自动双向的**，即：同样一个记忆库文件既可以用于英译汉中，也可以用在汉译英中。**词库是单向的**，例如：创建好了一个英汉词典并将它导出，那么这个英汉词典只能用在英译汉的翻译项目中；若需要用在汉译英项目中，则需要将词库进行反向，词库反向的方法：新建一个汉译英的翻译项目，在“项目词典”中导入刚才导出的那个英汉词库，软件就会自动做反向处理了。

## 8. 我有很多翻译资源，如何将它们转入到雪人中来用呢？

答：(1)若是以前利用 WORD、写字板等做的翻译文稿，中英文分开的，可以利用雪人双语对齐的“读入原文、读入译文”功能进行句级对齐，把它们做成记忆库。(2)若是中英混在一起无法划分的双语资料，则利用雪人双语对齐的“双语粘贴”功能，粘贴后雪人可以自动将混在一起的中英文分开，并进行句级对齐。(3)若是 TRADOS 等 CAT 软件直接生成的译稿文件，则利用雪人双语对齐的“导入双语”功能进行句级对齐，做成记忆库。(4)若是 XLS、XML、TXT、CSV 等格式词汇，则直接在雪人的“项目词典”中导入就行了。

## 9. 记忆库大了，软件使用起来会不会感觉较慢？

答：不会的。雪人的速度是所有 CAT 软件中最快的，在 32 位操作系统下，可以支持数百万句的记忆库，搜索速度每秒大于 50 万句。如果使用 64 位版本，超过 1000 万句也没有问题。

## 10. 记忆库的最小匹配度与片段搜索的区别？

答：记忆库的最小匹配度范围：0.5~1，软件在**项目设置**中默认设为 0.5，表示对于相似度在 0.5 以上的句子，会自动显示在屏幕下方的**记忆库**页签中，双击或按 F6 键可直接在译文中输入，并对不同的部分自动做了相应的替换。“片段搜索”是根据用户鼠标划选的部分词语从记忆库中查找符合的结果，具体操作方法：鼠标在原文处划选某片段后，在下方的**片段搜索**页签立即显示出搜索结果，若长句看不完整，可点击旁边的**T**按钮，转换到“文本方式”查看；按【Ctrl+C】、【Ctrl+V】可以将查找的结果复制到译文中或定义新词中。

## 11. 雪人与 trados 等的兼容性如何？

答：雪人与 trados、雅信等其他 CAT 软件有着良好的兼容性，可直接导入雅信的记忆库，或者将记忆库文件转成 TMX 格式导入。术语、词典可以先将它转换成 Excel、TXT、XML 格式的再导入。

## 12. 定义术语和定义新词的区别？

答：术语是在整个翻译项目中需要统一译法的词，一般情况下一个术语只对应于一种解释，普通的词汇是比较随意的，可以包括短句，例如把“long long ago”，“定于为“很久很久以前”，在下次碰到时，点击就可以输入了，节省了打字的时间。一般术语都是指专业的词汇，有些翻译项目，对方会提供术语库，这时，将它导入到**项目词典**中利用【**右键菜单→设为术语**】功能将它们设定为术语就行了。

## 13. 如何进行句子的拆分与合并？

答：雪人在导入原文时对句子进行了准确的拆分，同时亦有方便的合并与拆分功能。把光标移到需要拆分的地方，按【Enter】键，即可把当前的句子拆分为两个单元句子；把光标移到需要合并的第一句话的末尾，按【Delete】键就可合并两个单元句子。注意：拆分或合并句子功能仅限于在**对照模式**中，**单句模式**中不支持，需要切换回对照模式下进行。


## 14. 需要拷贝哪些文件就可以将工作移到其他机器上？

答：雪人将项目词典、项目记忆库及用户翻译的文件全部保存到后缀为 .STP 的文件中，只要拷贝这个 .STP 文件及用过引用添加的“词典、记忆库”文件（注意：词典、记忆库的



相对路径应该与原来的相同)。此外,雪人还自动生成了 STP.BAK 的备份文件,备份文件与.STP 的内容一样,一般不用理会它,但若.STP 意外损坏,只要将它的后缀名.BAK 这几个字去掉,就可以在雪人中打开它了。

## 15. 免费版词汇量有上限吗?

答:只要内存足够,没有具体上限,一般在 32 为操作系统下 500 万词条是可以支持的。在  **项目词典** 不要直接导入太大量的词汇(超过 50 万会影响机器的效率)。大型的词库应该尽可能的使用引用“添加”的方式,这样可以节省占用的资源。

## 16. 免费版没有词典和记忆库合并功能,该如何合并呢?

答:标准版有词典、记忆库合并功能,方便积累,而免费版没有这项功能。对免费版用户建议:建立一个专门的翻译项目用于积累词典和记忆库,通过导入、导出实现积累。具体做法: 1. 新建一个翻译项目; 2. 导入词典或记忆库。3. 导出雪人格式的词典或记忆库供其他翻译项目添加引用。

## 17. 如何维护雪人的项目记忆库、项目词典?

答:软件默认状态是长光标的浏览状态,在浏览状态不能修改,按【F2】可以进行编辑和浏览状态的切换,编辑状态下可以对译文、原文进行修改。在浏览状态下按【Ctrl+Delete】键可以删除整行记录。

## 18. 雪人正版软件能在几台机器上用?

答:同一时刻只能用在一台机器上,不同时使用可用于 3 台机器上,只要插入软件狗就行了。软件狗只记录 3 台机器的信息,超过 3 台的话,需要重装一下再用。

## 19. 雪人正版软件的升级是免费的吗?

答:(1).在同一个版本免费升级(版本号:V1.XX 内)

(2).购买后一年内如果发布新版本(版本号:V2.XX)可以申请免费升级。(注:“一年内”指以购买时的收据或发票的日期起算的一年内,并且在这一年内提交申请,逾期申请无效。)

20. 规则词典定义中,为什么“[ ]”这样的字符显示不出来?比如原文:[数字] 译文:[数字],规则定义成:原文:[{digits}] 译文:[@1],但是“[”这个字符,结果显示不出来?

答:因为软件在解释词义时候,[ ]包括部分被认为是备注,所以需要加上转义符号,正确的写法是:原文:[{digits}] 译文:\\[@1]。注意:译文需要同写成\\[@1],定义词语的词义也是一样,把词义定义为“[xxx]”的时候选词会选不出来,如果需要出现“[xxx]”

的话, 需要把词义写成“\[xxx]”, 其他括号例如: ()<>{} 等等的定义方法也类似。

## 21. 为何在 Vista 或 Windows 7 下无法保存设置状态?

答: 因为在 Windows 7 或 Vista 中, 需要管理员才能安装软件, 安装后安装目录的修改权限普通用户没有, 软件无法把配置文件保存在安装目录下。为了解决这个问题, 可以把安装目录的修改权限设置给当前用户。(在软件的安装文件夹中点击右键选择“属性”, 在“安全”页面中点击“高级”, 在对话框中加入当前用户的修改权限)

## 22. 为何免费版本的“在线词典”显示为空白?

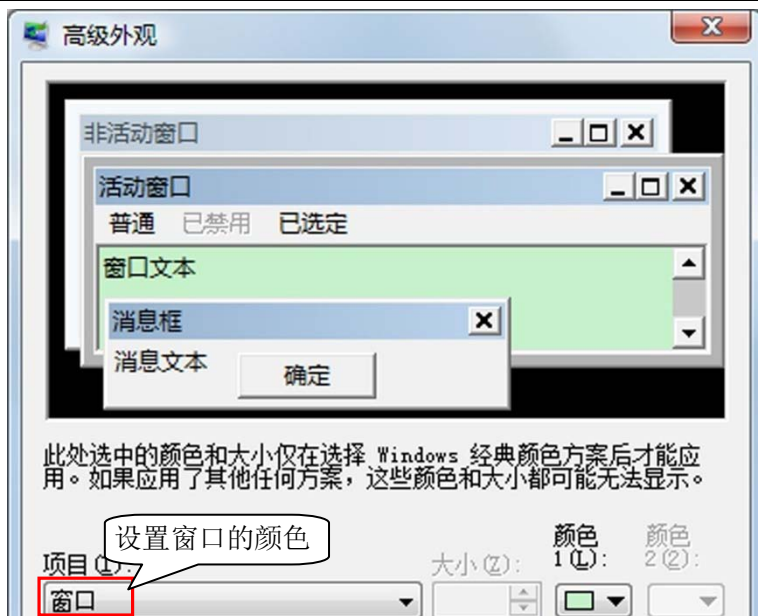
答: 主要有两种原因: 1. 没有在解压后的文件夹中运行或者是把运行程序直接拖放到了桌面中, 解决方法是在解压目录下运行或者在桌面上建立快捷方式; 2. 运行的操作系统不是中文版的操作系统, 解决方法是把解压后的目录名“雪人翻译软件”改为“SCAT”就可以了。

## 23. 免费版直接双击 STP 文件为何不能打开? 该如何关联?

答: 因为免费版是绿色版本, 无需安装, 解压即可运行, 所以没有建立打开雪人 STP 文件的关联软件(标准版不存在这个问题)。免费版的用户可以自己设置来解决, 方法: 1. 找到一个 STP 文件, 点击它, 然后点击【**右键菜单→打开方式**】2. 点击弹出窗口的“浏览”按钮, 选择雪人 CAT 软件, 并注意下方“始终使用选择的程序打开这种文件”前面的选项“打勾”, 就可以建立双击关联打开了。

## 24. 如何设置操作界面的背景色及字体大小?

答: 雪人 CAT 使用 Windows 的背景颜色作为背景, 只要设置 Windows 的背景颜色就可以了。如果使用的是 Windows XP, 在桌面属性中选择“外观”页, 点击“高级”, 选择“窗口”项目, 将“窗口”的背景颜色设为自己希望的颜色即可。若是 Vista 操作系统, 在桌面中右击鼠标, 选择【**右键菜单→个性化→WINDOWS 颜色和外观**】功能, 然后点击“高级”按钮, 设置“窗口”的颜色, 如下图:



设置字体大小方法：点击【工具→软件设置】菜单，在“软件设置”窗口中可以调节字体的大小。进行背景色及字体设置后，需要重新启动雪人CAT才能生效。如下图：



## 25. 导入、导出 WORD 文档时怎么软件好像没有反映？

答：软件在导入 WORD 时，需要对 WORD 文档中的字体、样式等等进行分析，需要花费一定的时间。导入、导出 WORD 的速度大约是每分钟 10 页，如果文件页数较多，可能需要比较长的时间。建议导入、导出前，关闭已打开的 WORD 窗口。不过导入是一次性的功夫，以后打开是很快的。

## 26. 导出 WORD 文档时出错，怎么办？

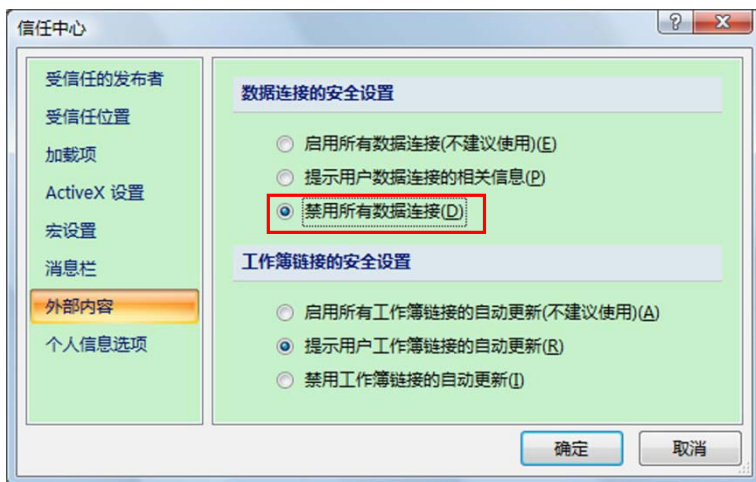
答：若 WORD 的原文设置了语言，当翻译为另一种语言时，通常会提出拼写出错；或如果文章比较长，WORD 的拼写检查会报出错，这样就导致了不能正常导出。解决办法：关闭 WORD 的拼写检查功能。若是 WORD2007，点击“圆按钮”下面的“WORD 选项”，点击“校对”来关闭 WORD 的拼写检查功能。若是 WORD2003，点击“工具”——“选项”——“拼写与语法”标签页。对于原文设置了语种的文件，在导出 WORD 后最好在 WORD 中全选修改为译文语种，这样 WORD 的显示就会正确很多。

## 27. 导入 EXCEL 文件时出错怎么办？

答：如果导入的 EXCEL 文档是写保护的，则不能正常导入，解决办法去掉 EXCEL 文档保护。

## 28. 导出 EXCEL 时为什么有时一直无法完成？

答：因为某些 EXCEL 文档中存在外部数据连接，在 EXCEL 中设置禁用所有的数据链接后，就可以正常导出了。EXCEL2007 具体设置方法：点击【开始】→【EXCEL 选项】→【信任中心】→【信任中心设置】→【外部内容】，选择“禁用所有数据连接”选项，然后再导出，如下图：



## 29. 使用 Office2007 怎样生成 Office2003 兼容的文档？

答：在 Office2007 中导出的译稿文件，若希望在 Office2003 版中亦能正常打开，则需要先在 Office2007 中打开此译稿文件，然后另存为 Office2003 兼容的文档。

## 30. PDF 转 WORD 一句插入了其他内容而被拆成多句的翻译？

答：将其他的句都集中在某一句来翻译，在其他句的相应译文处则不翻译，并输入 1~2 个空格。

## 31. 异地如何进行多人协同翻译？

答：使用雪人 CAT 网络协同翻译平台，可以实现局域网或互联网的多人协同翻译。实时共享记忆库和术语库；实现翻译与审校同步；平台整合即时通讯、文档管理、论坛于一体，方便各地译员、审校即时沟通。“雪人 CAT 协同翻译平台”由服务器端软件和客户端软件构成，雪人官网有绿色免费版下载。

## 32. 为什么出现保存的文件不能打开？

答：这种情况一般发生在用 U 盘存储的文件上，由于 U 盘存储文件容易出错，而且一旦出现 U 盘介质损坏，文件很难修复，有时连 .BAK 和 .SAV 的备份文件也一并损坏。用户直接在 U 盘上工作时，应防范这方面的风险，建议在硬盘上工作，把 U 盘作为文件的备份，这样可以更好地保护您的劳动成果。

## 33. 导入 PPT 文件出错时怎么办？

答：可能是 PPT 中包括特殊的内容，暂时的解决办法是把所有页面拷贝后粘贴在一个新 PPT 文件后就可以正常导入。

## 34. 为什么 PPT 组合内容没有导入？

答：PPT 中组合的内容如果没有取消组合，这部分内容目前版本不会导入，如果需要翻译，需要先取消组合后再导入（取消方法：在 PowerPoint 中点击【视图工具栏→绘图→取消组合对象】）。

## 35. 为什么 WORD 中带批注不能导出？

答：不能导出的原因是：原文带批注，并且当前用户无权修改这个文档时，所以导不出。解决的办法是：在 WORD 中将原文复制一份，并接受所有的批注，然后再导入雪人，使用记忆库预翻译快速翻译再导出。

### 36. 为什么 WORD 的文本框内容没有导入？

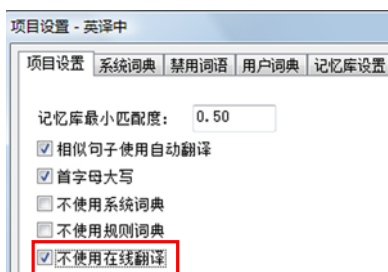
答：导入了。产生这种错觉的原因是文本框的内容不一定紧跟着正文部分，为了不影响正文，在“对照模式”下文本框的内容显示在文章的末尾。在“单句模式”下，可以在预览窗口中，**双击**文本框的内容，软件会自动跳转到相应的句子来翻译的。

### 37. 为什么 WORD2010 Starter 版不能正确导入？

答：因为 Starter 2010 是微软指定的 Microsoft Works 的替代产品，是一款低端产品，并不提供公开销售和下载，而是提供给 OEM 厂商预装在系统中。它并不能兼容 office 所有的文件格式，无法支持创建、编辑和运行宏命令等等许多功能被限制了。

### 38. 为什么“在线自动翻译”不能使用？

答：软件除了在使用“在线词典和在线翻译”时才向互联网发送请求翻译的内容，除此之外，不会向互联网发送任何数据。为了避免在线翻译意外发送了待翻译的句子，雪人 CAT 在 V1.21 版后，新建的项目文件默认不启用“在线翻译”功能，如果用户需要使用该功能，可以点击【项目管理】→【项目设置】设置为使用在线翻译。如右图：

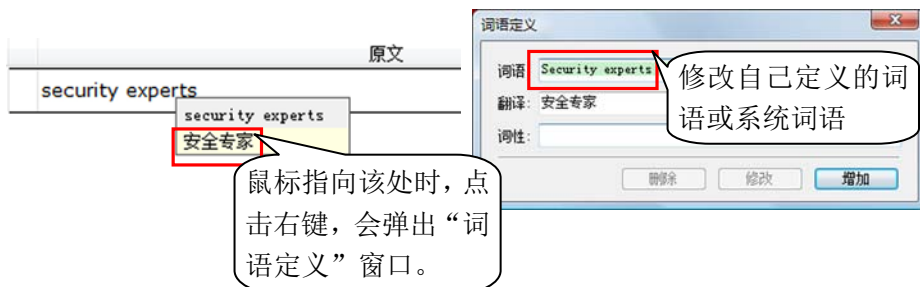


### 39. 导入原文时，显示转换器的问题，是怎么回事？

答：这种情况一般发生在机器上安装了 WPS 软件后，WPS 修改了 WORD 的一些默认设置而产生的，解决办法：重新安装 WORD。


### 40. 如何修改所定义的术语、词语？

答：在 **项目词典** 中按【F2】将光标切换到编辑状态后可以进行修改，或者点击【右键菜单→修改词语】功能。如果是翻译时遇见需要修改的词语，将鼠标指向该词时，在弹出的屏幕取词窗口中，鼠标指向具体的词义并点击右键，此时会弹出“词语定义”窗口供修改词义。如图：





## 41. 如何快速处理非译成分（或快速复制原文）？

答：整句复制原文按【F8】键。若需部分复制原文，鼠标划选后直接拖放到译文处或则按  按钮。若整句原文、译文都需要保持一样的，最快捷的处理方法是不翻译，将译文留空，软件在导出译文时，对未翻译的句子会自动以原文代替；不过，质量检查会提示这些未翻译的句子，解决办法是按【F9】将它设为审核正确的句子。还可以在“预翻译”窗口中的“非译成分”中预先定义不需要翻译的内容，这样预翻译时软件会在备注栏以“NTR”标注，并在该句前面打“√”，表示该句的译文完全正确。软件将认为该句使用原文进行翻译，质量检查时不会被认为是未翻译的句子。

## 温馨提示

由于U盘存储文件容易出错，而且一旦出现U盘介质损坏，文件很难修复，有时连.BAK和.SAV的备份文件也一并损坏。用户直接在U盘上工作时，应防范这方面的风险，建议在硬盘上工作，把U盘作为文件的备份，这样可以更好地保护您的劳动成果。