

机器翻译对俄语笔译的辅助作用

赵小玉

吉林大学外国语学院, 吉林 长春

收稿日期: 2024年3月26日; 录用日期: 2024年5月20日; 发布日期: 2024年5月31日

摘要

我们现在常用的机器翻译包括统计机器翻译和神经机器翻译, 更多的是两者并用的混合翻译, 利用大量的双语语料库和深度学习技术, 机器翻译兼具技术与人文的双重特点。机器翻译需要技术发展和丰富语料库的共同支撑, 我们对如何更好地使用机器翻译的探讨和实践, 可以提升工作效率以及翻译能力, 反过来推动更精准、高效的科技应用于机器翻译, 这印证了机器翻译与人工翻译相辅相成的优势。

关键词

ABBYY FineReader, Word, 机器翻译

The Auxiliary Role of Machine Translation in Russian Translation

Xiaoyu Zhao

College of Foreign Studies, Jilin University, Changchun Jilin

Received: Mar. 26th, 2024; accepted: May 20th, 2024; published: May 31st, 2024

Abstract

We now commonly use machine translation, including statistical machine translation and neural machine translation, and more often than not a hybrid translation using both, using a large number of bilingual corpus and deep learning technology, machine translation has both technical and humanistic characteristics. Machine translation requires the joint support of technological development and a rich corpus. Our exploration and practice of how to better use machine translation can enhance work efficiency as well as translation capability, which in turn promotes the application of more accurate and efficient technology to machine translation, confirming the complementary advantages of machine translation and human translation.

Keywords

ABBYY FineReader, Word, Machine Translation

Copyright © 2024 by author(s) and Hans Publishers Inc.

This work is licensed under the Creative Commons Attribution International License (CC BY 4.0).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

1. 引言

机器翻译是翻译学习中避不开的一个功能，机器翻译可快速大量地翻译语言文本，当然，也存在机器翻译出的内容与原文之义相差甚远的情况，所以较为完整地翻译一篇文本，需要机器翻译结合人工翻译的检验[1]。本文利用 ABBYY FineReader 和 Word 软件探讨如何更好地使用机器翻译，意在实现利用机器翻译辅助思考、辅助学习的目的。

2. ABBYY FineReader 识别技术的应用

我们常常会遇到需要提取 PDF 文档中内容的任务，但操作过程较为麻烦，由于时间和精力限制，我们一般不会选择逐字母输入识别。手机上有较为便捷的拍照识图提取文字功能，但其对中文的识别尚可，一旦识别俄语就会出现乱码。

在日常的信息化识别中，我们常使用到 OCR 技术，通过扫描仪或数码相机检查纸上打印的字符，并通过检测暗、亮的模式确定其形状，然后用字符识别方法将形状翻译成数字化的文档[2]。但我们在实践使用中发现，大多数的扫描识别都只能识别单一的语言，在至少两种语言的翻译文本中，我们使用这样的识别方式，只能识别出一种语言或者直接会出现乱码。如何解决多种语言的识别问题呢，在笔译实践课程中，我们初步尝试了 ABBYY FineReader 软件，主要利用该软件的 PDF 提取功能(以下简称 ABBYY)。

ABBYY 是一个强大的解决方案，方便用户查看、创建、注释、编辑、合并和保护 PDF 文档[3]。以俄语为例，在 ABBYY 中打开 PDF 文档，选择好 OCR 语言后即可识别，可以选择保存整篇识别好的文档，也可选择一部分复制粘贴。利用其中裁剪和擦除的功能可以只保留和识别需要的内容，不受边角、标题、页码的影响。大量处理语言文本的过程中总会出现一些纰漏，利用其中 OCR 语言识别并验证的功能可以直接在 ABBYY 中修改，软件会用突出显示标出它认为的可能识别的有误之处，我们可以在校对中修改并保存。经过一些文本的实践处理，我们可以发现 ABBYY 在识别俄语时出现的一些常见错误，比如：会把 **ы** 识别成 **61**、**6!**，把 **й** 识别成 **N**、**访**等，这些错误可以通过多次大量的实践积累修改经验。

利用 ABBYY 软件，我们可以较为便捷地提取一些电子书上的内容，经过提取的文档更加方便整理、修改和批注。

3. Word 中查找替换功能的应用

Word 是一款信息化的办公软件，可以方便快捷地处理文字信息。处理机器翻译后的文本离不开 Word 中查找替换的功能，如上文提到的，ABBYY 识别后的内容会出现一些常见的错误，而这些常见的错误可以通过查找和替换实现批量处理，省去了在 ABBYY 软件中一个一个修改的麻烦。我们使用俄汉的机器翻译，面临的一个重大问题就是文本的对齐问题，汉语的几个字可能对应俄语中一个加标点的很长的句子，在汉语的文段中可以直接全部删除空格，但俄语单词之间必然会有空格，那么当文档中是俄汉

混合的内容时，就不能直接删除整篇文档的空格。

利用查找替换功能可以很好地解决这些困扰[4]，比如：将汉字加前空格变成红色，汉字加后空格变成红色，红色空格变成无，本质上就是将汉字前后的空格替换成带有特殊标记的空格，之后把带有特殊标记的空格删掉，这样就解决了在俄汉混合的文档中只删除汉字之间空格的问题。再比如，从 ABBYY 中粘贴来的内容可能会有很多无用的段落标记，同样利用查找替换功能，将汉字段落标记汉字变成红色，再将红色的段落标记变成无，就可以解决去除段落标记的问题。

Word 中用于处理翻译文本的常用通配符[5]有^p 或^13 (段落标记)、^#(任意数字)、[一-龠] or [一-隲](中文汉字(CJK 统一字符))、[a-zA-Z](任意英文字母)等，利用查找替换功能还可以实现将字母与汉字分隔开，提取专业术语的功能，在此不作赘述。

4. 文本提取后以表格形式对照分析

在利用 ABBYY 软件提取文本内容、Word 软件整理对齐后，我们可以采用表格的方式进一步对内容进行分析、整理和思考(见表 1)，该操作可使用 Word 中文本转表格的功能(以下摘自新华社评论[6])。

Table 1. Comparison of differences in Chinese to Russian translations by different translation software

表 1. 不同翻译软件俄语译文对比

新华社原文	DEEPL 翻译	有道翻译	Yandex 翻译
节日的意义，不只在于致敬。	Значимость фестиваля не ограничивается только данью уважения.	Смысл праздника не только в том, чтобы отдать дань уважения.	Смысл фестиваля не только в том, чтобы отдать дань уважения.
迈向“十四五”，我们的任务更繁重、目标更远大，更需要广大妇女争当社会主义现代化国家的“建设者”，更需要广大妇女弘扬爱国情怀、促进家庭和陆、传承家庭美德，争当文明风尚的“倡导者”，更需要广大妇女在新时代的宽广舞台敢于追梦、努力圆梦，	В 14-й пятилетке наши задачи становятся более сложными, а цели—более амбициозными, и нам нужно, чтобы больше женщин стремились быть “строителями” современной социалистической страны, чтобы больше женщин несли патриотические чувства, способствовали семейной гармонии, передавали семейные добродетели и стремились быть “защитниками” цивилизованного стиля. Тем более необходимо, чтобы большинство женщин осмелились осуществить свои мечты и стремились воплотить их в жизнь на широкой сцене новой эпохи, создать лучшую жизнь своими руками и добиться прогресса в общем развитии страны,	К “14”, наш мисс больше бременительны, цел больше женщин соперничать социалистическую модернизацию стран “строитель”, нужен больше широкий женщин р азвива патриотическое чувство, содействуй семье, передавай добродетели, ставь Мод цивилизации “защитник”, нужен больше широкий женщин в новой эре больше осмел воплотить, у сил мечт мечт сцен,	В преддверии “14-й пятилетки” наши задачи становятся более обременительными, а наши цели—более амбициозными. Нам нужно, чтобы большинство женщин стремились быть “строителями” современной социалистической страны. Нам также нужно, чтобы большинство женщин поощряли патриотические чувства, способствовали семейной гармонии, наследовали семейные ценности. добродетели и стремились быть “защитниками” цивилизации и моды. Нам также нужно, чтобы большинство женщин осмелились преследовать свои мечты и упорно трудились, чтобы воплотить их в жизнь на широкой сцене новой эры,

续表

<p>凭借双手创造美好生活，在全面建设社会主义现代化国家的伟大事业中更好实现自身价值。</p>	<p>Тем более необходимо, чтобы женщины осмелились осуществлять свои мечты и стремились к их воплощению на широкой сцене новой эпохи, чтобы своими руками создавать лучшую жизнь и лучше реализовывать свои ценности в великом деле строительства современной социалистической страны.</p>	<p>Создавая лучшую жизнь своими руками, лучше реализовывать свои ценности в великом деле всестороннего строительства социалистической модернизации страны.</p>	<p>Создавайте лучшую жизнь обеими руками и лучше осознавайте собственную ценность в великом деле всестороннего построения современной социалистической страны.</p>
---	--	---	---

以汉译俄为例，我们选用了新华社一篇评论的部分材料，对比以上译文，我们可以发现，不同软件的机器翻译翻译出的篇幅长短不同，用词也有很大的差别。有的译文篇幅短简洁，但对于初学者有些难以理解，有的译文篇幅较长，一个词语需要几个相近的俄语词汇才能表达清楚。

翻译一篇文章既是在检验自己所学的知识 and 技巧，同时也是在实践中提升自己的翻译能力。利用多款翻译软件翻译内容的对比，我们可以在其中学到更多的词语、句法以及表达方式等，以上述表格为例，“爱国情怀”、“家庭美德”、“全面建设社会主义现代化国家”等常用词，各个翻译软件都用了不同的词和表达方式，而有一些地方用的是同一个词，比如，“倡导者”、“建设者”。有时候信息冗杂，我们在所学的大量词语中一时无法分辨哪些是常用词和常用表达，那么，采用这种表格对比的方式，我们可以更清晰地知道该用哪些表达，才能让我们自己翻译的文本更地道和通俗易懂。我们在翻译时可以选用两到三种不同软件的机器翻译，采用表格的形式，将不同的译文进行对比，也可再添一列进行自己的翻译，对比之下，择优而取。

ABBY 是一款便捷的识别和提取方案，但任何一款识别软件的识别程度都是有限的，使用过程中需要使用者本身具备较高的知识水平和辨别力，在多次使用和检验中提升文本的准确度。在使用 Word 处理翻译文本时，不必拘泥于以上方式，更多探索，更多思考，才能发现更多的方法来辅助学习。

5. 结语

随着科学技术和大数据的发展，海量的信息在不断产生，这给译者带来严肃考验的同时，也给机器辅助人工翻译提供了机遇，日新月异的信息资源既是译者需要翻译的内容，也是多语料库的有力支撑 [7]，为了提高译文的准确度和工作效率，我们需要借助机器翻译，而正确有效地使用机器翻译也是一个译者需要具备的素质。当然，如果过于频繁且不加思索地使用机器翻译，那么机器翻译就会成为积累知识、提高思考能力的最大阻碍，所以，我们需要有计划、有方法地使用机器翻译，在使用机器翻译时探索一些辅助翻译的软件，在实践中发现软件的优点与不足，在多个软件中取长补短，完成自己独特的翻译文本，有能力者甚至可以给软件的改进提出一些建议。

总之，机器翻译与人工翻译是相辅相成、相互促进的关系，机器翻译的水平依赖于人工的检验，只有译者的水平高于机器翻译的水平，才能去检验和校正机器翻译的文本。我们对于 ABBY 和 Word 软件的探索还有很大的空间，在翻译中思考，在翻译中积累，在翻译中提高能力，是我们使用机器翻译的初心。

参考文献

- [1] 万佳琦, 肖飞. ABBY Aligner 和 Tmxmall 对齐工具主要功能探究对比[J]. 英语广场, 2020(9): 27-29.

- [2] 王琴, 王宇春. 四款主流语料对齐工具性能对比探析[J]. 南阳理工学院学报, 2024, 16(1): 45-52.
- [3] 张秀常. 中英文混合识别的利器——ABBY FineReader [J]. 中国教育信息化, 2012(9): 73.
- [4] 王蓓. 巧用 Word 的“查找与替换”功能[J]. 信息与电脑(理论版), 2010(10): 175.
- [5] 张越. Word 查找与替换功能的高级应用[J]. 科技创新与应用, 2012(16): 34.
- [6] 新华社评论员: 凝聚砥砺奋进的“她”力量——写在“三八”国际劳动妇女节[EB/OL]. 新华社. https://www.gov.cn/xinwen/2021-03/09/content_5591664.htm, 2021-03-09.
- [7] 周秦超, 林向阳. 混合机器翻译的研究现状和发展趋势[J]. 产业观察, 2024(1): 151-153.