

《伤寒论》人机翻译对比研究

张凯茜 赵旭

(陕西国际商贸学院 712046)

摘要: 基于神经网络的机器翻译使得译文的流畅度和准确性得到了提升, 但机器翻译仍存在局限, 选词、语序和内涵意义表达都有限制。学界研究关注机器翻译在正式语篇翻译实践与和译后编辑模式的探究, 对于文言文人机对比翻译研究的较少。本文以《伤寒论》为例, 通过对比人工翻译和机器翻译译文, 从通过数据进行可读性分析, 并从衔接和连贯角度对语篇进行分析, 探究机器翻译在中医典籍英译中的特征和不足, 促进机器翻译的不断完善和发展。

关键词: 人机翻译, 机器翻译, 《伤寒论》, 英译对比研究

1. 引言

随着现代技术不断进步, 机器翻译在几十年中飞速发展, 从初期基于规则, 过渡到基于统计, 到如今基于神经网络的方法, 高速的信息处理和加工为翻译实践活动提供便利。机器翻译也常应用在科技、政治、新闻等正式语篇中, 对于古汉语的翻译研究相对较少, 尤其是中医典籍。本文选择《伤寒论》为原材料, 语言简单、直白, 却省略多、内含隐形的逻辑关系, 给机器翻译提供了可能和挑战。通过对百度翻译、搜狗翻译和谷歌翻译的一本, 笔者选择了译文相对准确的谷歌翻译(简称谷译), 人工译本选择2017年出版的李照国译本《On Cold Damage》(简称李译), 通过数据分析和译文分析进行人机对比研究。

2. 人机翻译数据分析

关于译文的语法层面的准确性和可读性, 采用 Grammarly 在线语法纠正和校对工具进行分析, Grammarly 能基于云计算技术在英语语境下纠正语法错误、检查单词拼写、标点符号和句式结构并给出建议^[1]。检测的样本为谷译的宋本《伤寒论》, 检测结果: 整体得分: 67分(总分100分), 基于 Flechreading-easetest 可读性得分74分, 大多数成年人是可以读懂的, 因为谷译在选词上简单易懂; 李译可读性评分是58分, 对读者的英文水平有一定要求(可读性60-70分才能被80%的英文读者轻易接受)。

笔者通过语言数据网(languagedata)对李译和谷译的阅读难度进行测试, 对译文阅读难度等级进行测试^[2], 测试结果如表1。

类别	李译		谷译	
	量化指标	对应等级	量化指标	对应等级
词汇难度(VC值)	6.3	六级	5.11	四级
句法难度(SC值)	5.99	四级	4.76	高考
文本难度(LD值)	6.64	六级	5.21	四级

表1 阅读分级指南针

结果显示, 李译从词汇、句法和文本难度都高于谷译。通常情况下, 文本难度与文本可读性成负相关, 文本难度越高, 可读性就越低^[3]。《伤寒论》中有较多文化负载词和专业术语, 既具有文学性, 又具有科学性。读者群体主要是国内外中医学者和喜爱中国文化的读者。因此, 译文的可读性除了字句的难易, 还要看译文的表达是否准确。

3. 人机翻译词性构成

为了对比人工翻译和机器翻译在选词上的区别, 笔者将《伤寒论》原文、谷译和李译文分别整理成语料库, 从中选出20条通过 Readable 在线平台进行词性对比分析, 结果如表2。

词性	李译(比重)	谷译(比重)	词性	李译(比重)	谷译(比重)
形容词	9%	8%	名词	31%	28%
副词	2%	1%	专有名词	1%	1%
连词	10%	12%	介词	13%	10%

限定词	9%	11%	代词	3%	6%
感叹词	0%	0%	动词	21%	22%

表2 谷译和李译的词性分布比重

通过数据分析, 可以看出机器翻译的特征。与机器翻译相对较, 李译中使用名词和介词比重更大, 机器翻译中使用连词、限定词和代词比重更大。副词、专有名词和动词的使用比重差别不大。从文本数据分析看, 李译单词和句子数高于谷译, 但是每句单词数相差不多, 李译更常用被动句式, 更偏向于科技英语。科技英语的特点是被动句、多非谓语动词形式、名词化的, 译文严谨且衔接紧密。

4. 人机翻译衔接和连贯分析

通过测试可读性和文本数据分析后, 从衔接和连贯探讨译文的可读性。系统功能语言学家韩礼德(Halliday)和哈桑(Hasan)首次提出了“语篇衔接(cohesion)”的概念, 有五种衔接方法: 照应(reference)、替代(substitution)、省略(ellipsis)、连接(conjunction)及词义连接(lexical cohesion)。衔接侧重于对语篇内部的关系, 连贯则是对语篇意义的关联性的分析, 两者相互关联、相辅相成、表里相依, 不可分离^[23]。本文根据数据分析结果, 机器翻译使用了更多的连词以及动词使用比重相似, 主要讨论两个方面: 一是连接词的使用差异。二是谓语的选择类型上的差异。

4.1 连接词

原文: 太阳之为病^①, 脉浮, 头项强痛而恶寒^②。

谷歌: The sun is a disease^①, the pulse is floating^②, the head and neck are sore and aversion to cold^③.

李译: Taiyang disease [is characterized by] floating pulse, stiffness and pain of the head and nape, and aversion to cold.

太阳病的病理表现是: 脉浮、头痛、项部强痛、怕冷。原句^②是解释说明^①关系。谷歌翻译中和李译翻译都多次使用 and。谷译可拆成3个小句, 从语法角度看: 3者缺失了并列连词 and。从连贯性看: 小句中, head 和 neck 用法正确, 但 andaversion 并列对象还是 headandneck, 与原文相左, 句义加工理解错误, 连贯性欠佳。值得肯定的是, 机器翻译按照英文文法的习惯, 增加了系动词, 将意合的中文转换成形合的英文, 但是机器翻译无法细分小句子之间的关系。

原文: 太阳病, 头痛至七日已上自愈者, 以行其经尽故也。

谷歌: Sun disease, the headache has healed on its own within seven days^①, and it is because of the exercise of its classics^②.

李译: Taiyang disease with headache that heals automatically over seven days is due to [the fact that it] transmits all through the meridian.

谷译使用连接词 and 将原文的语义分成并列的2小句, 李译并没有使用并列句的形式, 而是采用两个 that 引导的定语从句。谷译中的“anditis”没有语法意义和实际意义, 是赘述表达, 可以省略的。除此之外, “经”并不是经典的意思, 是“经脉”的意思, 谷译在词汇选择时因语料库内容的有限性, 选择了错误的表达。

(下转第232页)

出来,从事茶叶等农副产品生产经营、交通运输,二、三产业收入比重大幅度提升,农民人均纯收入突破1万元。在此发展历程中,屏口各族群众发扬自力更生、艰苦奋斗的作风,曾连续三年调集全村劳动力修通了九道水-屏口-黄连溪-芭蕉的循环公路,吴家坡、骡马洞、后山等一批通组公路,通达宣恩县庆阳坝、本市盛家坝镇石门坝的经济断头路,形成道辐射全村、纵横交错的公路网。

(四)社会主义新农村建设。2010年,屏口村作为“仙居恩施”恩施玉露生态旅游走廊上一处最具茶乡少数民族村寨风韵的节点,依托丰富的生态资源和民族文化资源,坚持走农旅融合特色道路,制定了《屏口村中心村建设规划》和《恩施市芭蕉侗族乡屏口村示范区提升规划》,通过整体规划、项目支持,落实基础设施建设资金1300余万元、特色民居建设资金700余万元,动员和组织农民自主投入2100余万元,按照“青瓦屋面,飞檐翘角,木门木窗,白脊白墙,咖啡群墙”的侗乡民居建设风格,新建和改造特色民居890户,实施“一池三改”建设323户,在河道、公路沿线和农户庭院栽植柳树、楠竹、桂花树等绿化林木2万余株,整修河堤2000余米,新建侗族风格的风雨桥4座、观景亭3座,建成九道水-屏口生态旅游公路12公里,筒车坝、唐家院子等村落成为新农村建设亮点和游客游玩观赏的重要景点。各民族和睦相处,践行社会主义核心价值观,引导居民注重公共环境,推崇健康文明新风尚,促进民族团结进步。打造“仙居人家”农家乐、农家宾馆8家,每年接待游客20余万人,旅游综合收入800余万元。2015年4月,屏口村党支部书记肖肖云获评“全国劳动模范”,赴北京参加了全国庆祝“五一”国际劳动节暨表彰全国劳动模范和先进工作者大会。

近年来,屏口各族群众以精准扶贫为重心,实施乡村振兴战略,推动特色产业发展,经济提升取得了长足进展。2022年,屏口村将实施彩虹山组传统村落集中连片保护利用项目,对8栋文物保护单位进行修缮,建设1座风雨桥,完善挡土墙、游步道、标识标牌等设施。村里发展了,村民富裕了,乡村面貌变美了,各族群众发展的劲头更足了。

结论。屏口民族发展史就是各民族交往交流交融的历史,就是各族人民共同团结奋斗、共同繁荣发展的历史。各族群众在屏口同呼吸、共命运、心连心的奋斗历程,铸牢了中华民族共同体意识,形成了平等团结互助和谐的社会主义民族关系。只要我们深刻认识铸牢中华民族共同体意识的重大意义,推动各民族对伟大祖国、中华民族、中华文化、中国共产党、中国特色社会主义的高度认同,促进各族群众在中华民族大家庭中像石榴籽一样紧紧抱在一起,手足相亲、守望相助,就能实现中华民族复兴伟大梦想。

时光荏苒,万山从中的传统村落和文化遗迹终将改变模样。回望历史,屏口发展的脚步从未停止。展望未来,随着乡村振兴战略的深入实施,屏口各族群众必将迈着坚实的步伐,与全国各族人民一同为中华民族伟大复兴砥砺前行,开启全面建设社会主义现代化国家新征程。屏口村的明天将会更加美好。(作者:李小雪,中南民族大学在读博士研究生湖北民族大学硕士研究生;廖建洲,恩施州武陵山经济研究会常务副会长、秘书长、湖北民族大学经济与管理学院客座教授;杨光,恩施州武陵山经济研究会理事)

注:本文系2022年度湖北省恩施州民宗委各民族交往交流交融历史研究课题成果之一)

(上接第214页)

4.2 谓语

除了韩礼德提出的衔接方法,句子衔接主要以主谓一致关系为框架,以时、体、态、性、数、格以及其他功能词为辅助的语法体系^[9]。在中英互译过程中,如何将中文中的多个动词进行转换,也是难点。

原文:太阳病,发热,汗出,恶风,脉缓者,名为中风。

谷歌: Taiyang disease, fever, sweating, aversion to wind, and slow pulse are called stroke.

李译: Taiyang disease, [characterized by] fever, sweating, aversion to wind and moderate pulse, is called wind stroke.

病理表现以4个动词短语并列,省略了“和”的含义。谷歌翻译将中文的并列成功转译成 and 连接的名词/名词短语/动词,但因为没有理清太阳病和病理症状的关系,谷译统一认为是主语。原文是主动语态,谷译和李译都译为被动语态,符合英文的习惯。但是在助词的选择上,谷译因为主语识别错误,用了复数 are 而不是 is。

原文:太阳病,发热而渴,不恶寒者,为温病。

谷歌: Sun disease, fever and thirst, not aversion to cold, is a warm disease.

李译: Taiyang disease, [characterized by] fever, thirst and no aversion to cold, is called warm disease.

“名为”和“为”意思一样,都是“称为”,李译对于类似的翻译都采用被动语态,而谷译则“名为”是被动,“为”则是主动。但是在此例中,句子的谓语动词是 is,是正确的。笔者认为并不是因为谷歌翻译正确识别了主语是 Taiyangdisease,基于之前的错误类型,谷译将 notaversiontocold 当成了主语。

5. 结语

谷歌翻译的《伤寒论》从 Grammarly 和 LanguageData 数据分析

上说,译文单词和句子简单,可读性高于李照国译本,但是文本内容分析时发现,谷译存在误译,读者会不知其所云,实际可读性不高。Readable 对比人机翻译的词性选择发现机器翻译多使用代词和连词。机器翻译能翻译语言的概念意义,对于文化和交际意义会无法编码,译文会出现连接错误,指代错误,用词不准确,语句不连贯等问题。然而,值得肯定的是,机器翻译具有稳定性,会进行谓语句形态的转换、简单的隐藏含义的再现等。人工翻译依赖于译者的语言能力和翻译环境,不确定性大,不同的译者侧重点不同。两者有不同的特点和规律,可以在了解机器翻译的特性后,根据文本特征选择机器翻译、译后编辑还是人工翻译。

参考文献:

- [1]段海鹏,龙晓明.博物馆文本英译可接受性研究[J].桂林师範高等专科学校学报,2022,36(02)
 - [2]王洪亮,绪可望.语篇衔接与连贯认知分析理论及范式问题探究[J].东北师大学报(哲学社会科学版),2021(05)
 - [3]曹德和,王卫兵.亦谈“衔接”和“连贯”的概念界定及其关系说明[J].北华大学学报(社会科学版),2020,21(06)
 - [4]魏光.人工翻译与机器翻译译文编辑比较研究[J].海外英语,2021(19):3.
 - [5]金檀,李世炜,刘丰恺.(2021).学术英语阅读素材的智能选取与教学应用研究.《外语界》,(1),20-27.
- 基金项目:2022年度西安市社会科学规划基金项目“中医药古籍医籍机器翻译可接受性研究”(22YZ08)部分研究成果。
- 作者简介:张凯茜(1993.10~),女,陕西西安,陕西国际商贸学院,硕士研究生,研究方向:英语笔译;
- 赵旭(1983.6~),男,陕西汉中,陕西国际商贸学院,硕士研究生,研究方向:中医翻译。