

科普类文本在线翻译质量研究：英汉对比的视角

林倬如 盛颖睿 崔晓宁 梁 硕

(华东理工大学, 上海 徐汇 200237)

摘要：“新文科”建设理念为理工类大学英语课程和理工类专业英语课程提出改革新课题。本文以科普类文本为语料，从文本特点和英汉对比视角出发，先设计科普文本在线翻译译文质量评价标准和规则，再以语段为评价单位制作译文评测问卷并进行问卷调查，统计数据，分析四个在线翻译平台翻译科普文本的译文表现，评估译文的优势与不足，筛选译文质量较优的在线翻译平台以帮助读者阅读英语科普文本。研究发现，各平台译文质量从高到低依次为 DeepL、有道、搜狗、百度，DeepL 在帮助英语科普类文本阅读时效果更为理想。

关键词：科普类文本；译文质量；在线翻译；英汉对比

一、引言

“新文科”建设理念为理工类大学英语课程和理工类专业英语课程提出改革新课题。英语科普类文本阅读促进理工类专业学习者拓展学术视野、学习学术英语写作，是内容优质的学习材料和教学材料。然而，科普类文本有其特殊的文本特点，例如专业术语丰富、长难句多且较为复杂，使其成为阅读学习的重点和难点。在线翻译工具能较好地帮助学习者学习英语科普类文本，然而，在线翻译工具种类繁多，学习者有时难以选择译文质量较好的在线翻译工具。虽然已有许多对在线翻译工具译文质量进行的研究，但是针对科普类文本和从英汉对比视角进行的研究较为有限。

本文以美国知名科普杂志 Scientific American 中的文本语段为语料，结合现有研究和文本细读制定译文评价标准，对有道翻译、DeepL 翻译器、搜狗翻译和百度翻译（下文简称为“有道”“DeepL”“搜狗”“百度”）四个在线翻译工具译文质量进行研究。研究过程为：1）依据已有文献，从科普类文本中兼有的通用文本翻译问题、科普类文本特有的翻译问题和英汉对比视角下的翻译问题出发，分析在线翻译平台在翻译科普类文本时存在的问题；2）制定科普类文本在线译文质量评价标准和评分规则；3）从 Scientific American 杂志中随机选取语段，再筛选四种在线翻译工具译文有差异的语段为语料，制作译文评测问卷，根据所制定标准对译文进行评分。

二、科普文本在线翻译译文质量研究现状

目前，科普类文本阅读需求日益增长、在线翻译工具众多、在线翻译技术迭代迅速，国内外不少学者对科普类文本及在线翻译译文质量评估展开了研究。

科普类文本有其特有的文本特点及翻译难点。方梦之（1999：32-26）最早将科技英语分为科普文章、科学论文和技术文本。徐彬和郭红梅（2012：42-46）否认科普英语带有科技英语的词义多、非谓语句多等显著特点，认为其与普通英语差别只在于其内容关于科学，科学技术词汇更多。以往研究对于科普类文本定义的界定可分为两类：一类观点认为，兼有两种文本的特点；另一类观点认为，科普类文本与通用文本类似，区别仅在于术语多少。在阅读英文原版科普类杂志后，结合已有研究，本文认为，科普类文本是介于科学类和文学类文本之间的一种特殊文体，兼具科学性和文学性。其科学性集中体现在科普类文本“术语多”，此外，还体现在长难句多这一方面。

90年代起，国内出现从某一英汉对比角度研究科普类文本翻译难点的研究。丁棣（1993：25-27）指出，英文“多词一义”的

现象会对科普翻译带来一定挑战；李丙午、燕静敏（2002：5-7）关注到科技英文中为提高语言客观性程度多用“名词化结构”；由此可知，采用英汉对比的视角来挖掘科普类文本翻译难点有研究基础，但目前还并未有研究者从英汉对比的视角多方面地研究译文质量评价。

随着在线翻译技术的发展，在线翻译译文质量成为新的研究领域。张霄军（2007：80-84）总结了翻译质量量化评价研究的两种主要途径：人工评测与自动评测。聚焦人工评价的理论模型，张卫晴和张政（2006：13-17）尝试将模糊综合评判方法应用于机器翻译的译文质量评价，建立“新的定量计算机翻译系统评价体系”。

综上所述，不少学者都对科普类文本及在线译文质量展开了研究，对在线译文质量评价的研究已有基础。然而，现有研究所制定的译文质量评价标准存在准则笼统、权重分配主观等不足之处。第一，一些对在线翻译工具进行人工评测的研究从“信达雅”和“译文忠实度、术语准确度、行文流畅度”等角度进行评测，在实际操作中极为依赖评者的主观判断，指标无法量化或有较大的随意性。第二，不同因素对译文质量有不同的影响，以均分权重的方式制定评价标准较缺乏科学性，部分文献中不同指标的分值占比存在差异权衡，但并未阐释确定权重的理论依据。

三、科普文本翻译问题类型

科普类文本特点及翻译难点是科普类文本在线翻译译文质量评估标准制定的基础。本研究依据已有文献，分别从科普类文本中兼有的通用文本翻译问题、科普类文本特点、英汉对比视角分析不同在线翻译平台在翻译科普类文本时存在的问题。

（一）科普类文本中兼有的通用文本翻译问题

科普文体是文学和科学相结合的写作体裁，此外，一般还具有通俗性和趣味性的特点（王振平，2006：36）。英译汉的译文评价也适用于科普类文本在线翻译译文评价。

1. 词义

词义选择对任何形式的翻译都是重点。从语言单位角度看，学术汉语词汇受语素和语块制约，相较于英文，汉语词义独立性较大，伸缩性较小。英语词汇含多层次意思，灵活性强，易产生误解，依赖于语言系统进行择义（蔡力坚，2012：78-82）。

2. 惯用表达

惯用表达背后有特定历史文化和社会因素。在翻译过程中，直译和简单的意译无法传达出看似简短的惯用表达背后所蕴含的信息。部分英语文本中的习语在汉语中有对应，部分则依赖于巧

妙的翻译,有时需结合语境选择惯用表达含义,在部分句子中对机器翻译产生挑战。

3. 句意消歧

歧义句是指句子会产生两种或两种以上的理解。在英译汉时,应识别句子逻辑、避免歧义发生。

4. 语篇连贯

相同术语或内容在一篇文章中前后不一致,不仅会迷惑读者,还会降低信息的可靠性与专业性。术语的不一致可能让读者认为涉及另一事物,使得语篇连贯性不佳。

5. 语言流畅

对语言流畅的考察是所有译文质量评价都涉及的方面。语言不流畅表现在译文生硬难懂、过于冗长等方面。

(二) 科普类文本特有的翻译问题

已有研究对科普类文本特点的讨论结果各有不同,但对科普类文本术语多这一特点认识一致。此外,涉及科技的英语文本中还出现长句来做客观准确的描述(余高峰,2012:182-186),考察在线翻译工具在长难句翻译上的表现也是重要方面之一。

1. 术语/专有名词/新词

科普类文本多涉及各领域专业名词和术语。英汉语中专有名词的定名都包含着复杂的思维活动,常采用概念命名、引申义命名、比喻义命名、形象特征命名等方法(仇伟,2015:13-16)。在进行专有名词翻译时需结合语境,选择直译、意译、音译,以达到形象准确传达原文的目的。

2. 长难句

科普类文本中多利用长句、复合句来分析数据和阐述事件过程、研究内容,用长句来做客观准确的描述(余高峰,2012:182-186)。此类长句结构复杂,通常包含几个并列的从句,从属的分句和条件假设句,各分句间环环相扣,关系交错。

(三) 英汉对比视角下的翻译问题

翻译不仅是语言的问题,还受到文化的制约和影响(连淑能,2010:16)。总体来说,西方人惯用理性思维,在英文表达中更注重严谨性和规范性;中文则体现出中华民族的文化传统,重感性、重表意。研究筛选英汉对比视角中对在线翻译工具译文质量产生较为明显影响的几个方面,最终选取以下四项作为评价标准。

1. 形合与意合

“英语是重形式的语法型语言。”(连淑能,2010:74)英文注重显性连接,以形显义。而“词语之间的语法关系主要是通过安排语序和使用助词来表达的。”(连淑能,2010:93)有时对原文进行逐词翻译会出现生硬情况,就要考虑将汉语“意合”的表述方式运用到译文产出中。

2. 静态与动态

“英语倾向于多用名词,因而叙述呈静态,汉语倾向于多用动词,因而叙述呈动态。”(连淑能,2010:104)名词化现象在英语科普类文本中更为常见,而在中文表达中,动词是表意的重要载体。

3. 替代与重复

英文表达追求简明,当需要重复提及同一个事物时,往往采用代词、修辞、抽象化等手法来形成指代关系;在中文里,“重复”作为一种修辞手法起到强调内容、美化结构的效果,符合中文重表意以及对文字美感的追求。在英汉翻译中,需注意到替代与重

复的文化差异。

4. 主动与被动

被动语态的使用也是英文表述的一大特点,尤其体现在科普类文本中。使用被动句,能在一定程度上消除主观臆断,使文章更显客观严谨。而中文较少使用被动表述。

四、科普文本在线翻译译文质量评价标准

依据上文对在线翻译工具翻译科普类文本时可能对在线翻译工具构成挑战的问题进行的分析,可以得到各评价指标,分别为科普类文本视角的词汇、惯用表达、句意消歧、语言流畅、语篇连贯、术语/专有名词/新词、长难句和英汉对比视角的形合与意合、静态与动态、替代与重复、主动与被动。然而,译文评价指标标准设计和各评价指标权重值的确定需要科学的依据。

(一) 译文评价指标标准设计

范守义(1987:2-9)提出用模糊数学的理论和方法对译文定量评价,从不同“方面”对译文以句为单位进行矩阵评价,每句只从一个“方面”进行观察,据此给出这个句子达到这个“方面”的隶属度,再将所有句子所得的隶属度相加并求其平均数,即为对该段译文所做出的评价。本研究拟采用模糊综合评价法建立译文质量评价标准。不同的是,评价以语段作为单位,用扣分法考察语段各方面的质量,再综合计算得分。

(二) 译文评价指标权重值的计算

各评价标准的权重确定一直是译文质量评估体系中的难点,难以量化和过分主观的问题将降低质量评价体系的可靠性和科学性。

Satty提出了1~9标度法,用区间数来描述评估指标两两比较的相对重要程度,用 a_{ij} 来表示 A_i 与 A_j 的重要性之比(孙逸群,周敏康,2017:20-24)。例如, A_j 与 A_i 重要性之比为 $a_{ij} = \frac{1}{a_{ji}}$ 。列出比较矩阵A,

$$A = \begin{bmatrix} a_{11} & a_{12} & a_{13} & a_{14} & a_{15} & a_{16} & a_{17} \\ a_{21} & a_{22} & a_{23} & a_{24} & a_{25} & a_{26} & a_{27} \\ a_{31} & a_{32} & a_{33} & a_{34} & a_{35} & a_{36} & a_{37} \\ a_{41} & a_{42} & a_{43} & a_{44} & a_{45} & a_{46} & a_{47} \\ a_{51} & a_{52} & a_{53} & a_{54} & a_{55} & a_{56} & a_{57} \\ a_{61} & a_{62} & a_{63} & a_{64} & a_{65} & a_{66} & a_{67} \\ a_{71} & a_{72} & a_{73} & a_{74} & a_{75} & a_{76} & a_{77} \end{bmatrix}$$

将A中的每行相加,进行归一化处理得到A的特征向量 W_1 。

本研究采用层次分析法制定评价体系中的标准权重,从科普类文本和英汉对比两个视角建立译文质量评价体系。科普类文本视角层面A中包括词汇 A_1 、惯用表达 A_2 、句意消歧 A_3 、语言流畅 A_4 、语篇连贯 A_5 、术语/专有名词/新词 A_6 、长难句 A_7 这7个方面。英汉对比视角层面B包括形合和意合 B_1 、静态和动态 B_2 、替代和重复 B_3 、主动和被动 B_4 这4个方面。其中A层面注重译文中的错误;B层面注重考察译文是否关注到英汉语言差异乃至其背后存在的中西方文化差异。从阅读感受的角度而言,注重英汉对比的视角能够减少由于译文“翻译腔浓”“语言生硬”给读者带来的阅读障碍,但B层面所强调的差异主要体现在英汉语内在逻辑思维的不同,对文意理解的影响不及A层面中所出现错误影响大,因此考虑A层面与B层面整体对于译文质量的影响,本文在做权重设置时,设置A, B=(0.7, 0.3)。

结合层次分析法与1~9标度法,计算得出A的特征向量

$W_1, W_1 = (0.12, 0.05, 0.22, 0.08, 0.06, 0.20, 0.27)$ 。同理可得 B 的特征向量 $W_2, W_2 = (0.35, 0.23, 0.19, 0.19)$ 。

(三) 译文评价指标标准

最终得到译文质量评价标准如下表所示:

表 1

		第一档		第二档	
科普类文本中兼有的通用文本翻译问题	词义 0.12	有少许词义翻译错误,但影响较小(扣1~4)	错误翻译词义,且影响较为严重(扣5~8)		
	惯用表达 0.05	注意到惯用表达,但翻译不准确,对理解影响较小(扣1~2)	错误翻译惯用表达/直译,严重影响理解(扣3~4)		
	句意消歧 0.22	语义存在部分歧义,但读者能通过上下文或文化背景理解(扣1~7)	语义存在严重歧义,表意不清,严重影响理解(扣8~15)		
	语言流畅 0.08	语言不够流畅,语序不符合汉语逻辑,但对理解影响较小(扣1~3)	语言生硬,逻辑混乱,严重影响理解(扣4~6)		
	语篇连贯 0.06	衔接词使用不准确,未保持术语前后一致,但对理解影响较小(扣1~2)	衔接词使用不准确,前后术语不一致,导致表意与原文偏差,严重影响阅读理解(扣3~4)		
	术语/专有名词/新词 0.20	直译/意译术语,与领域统一规定对不一致,但对理解影响较小(扣1~7)	直译/意译/漏译术语,导致译文与原文偏差,严重影响阅读(扣8~14)		
长难句翻译问题 0.27	未根据长难句逻辑进行语序重排,但对理解的影响较小(扣1~9)	未识别长难句内在逻辑,导致译文生硬难懂,严重影响阅读(扣10~18)			
英汉对比视角	形合与意合 0.35	未注意英汉明显差别,调整主谓顺序和结构且影响理解(扣1~5)	未将英语中连接词、逻辑语做适当删减和语序调整,严重影响理解(扣6~11)		
	静态与动态 0.23	未将英语中多名词现象做动态化处理,不符合汉语动态倾向,但对理解影响不大(扣1~3)	未将英语中多名词现象做动态化处理,不符合汉语动态倾向,且严重影响理解(扣4~7)		
	替代与重复 0.19	未将英语中的替代倾向转化成汉语中的重复倾向,但对理解影响不大(扣1~3)	未将英语中的替代倾向转化成汉语中的重复倾向,且严重影响理解(扣4~6)		
	主动与被动 0.23	未注意英汉明显差别,调整主谓顺序和结构且影响理解(扣1~5)	未将英语中的被动形式转化成汉语中习惯的主动句式,且严重影响理解(扣4~7)		

五、科普类文本在线翻译译文质量评价与分析

(一) 评价方法

本研究从美国知名科普杂志 Scientific American 中随机选取 10 个在线译文有较大差异的语段,制作调查问卷。每个语段约 100 词左右,内容涵盖神经科学、心理学、医学、生物学等多个领域。

问卷制作完成后,请 20 位研究英汉对比的翻译专业研二在读学生依据制定的译文评分标准对各个平台的在线译文进行评分。评分时,要求评分者批注评分依据。若同一错误在语段中重复出现,扣满 3 次为止。若同时出现英汉对比视角分项与长难句问题时,在英汉对比视角项下扣分。

统计数据,得出各在线翻译平台的平均分,以确定各平台的综合表现,及各个评价指标的扣分情况,以对在线译文质量进行详细分析。

(二) 统计与分析

发放问卷 20 份,回收有效问卷 20 份,得到评分情况如下表

所示:

表 2 在线翻译质量评估综合得分统计表

	语段 1	语段 2	语段 3	语段 4	语段 5	语段 6	语段 7	语段 8	语段 9	语段 10	均分
搜狗	91.4	92.0	84.8	91.6	89.0	82.5	82.4	85.8	80.2	85.5	86.5
百度	84.8	85.1	84.5	91.4	84.6	80.3	82.2	81.4	79.8	82.3	83.6
DeepL	89.7	94.0	88.4	93.1	94.1	93.2	86.9	91.8	86.8	83.0	90.1
有道	89.3	89.9	92.3	92.1	90.0	84.4	85.1	84.1	86.6	80.7	87.4

综合来看,DeepL 平均得分最高,译文的综合水平最高,其次是有道、搜狗,其分数差别较小,百度的扣分最多,译文的综合表现相对不佳。

统计问卷中每位评分员对每句在线译文在各指标上的扣分情况,取平均数,得到四种在线翻译引擎在各指标上的表现,如下图所示。扣分值大小表示译文质量的高低,扣分值越低的平台译文质量越高。

研究发现:

从各指标对比来看,各翻译器在词义准确性上的扣分最多,句意消歧、语言流畅、术语表达、长难句翻译等方面错误的影响占比类似,在惯用表达、语篇连贯、形合与意合方面占比类似,在静态与动态、替代与重复、主动与被动等方面错占比类似且最小。

六、结语

依据前人文献和译文研究制定的科普文本质量评价标准,结合模糊综合评价法和层次分析法确定更具客观性的标准权重。根据标准对所选译文进行质量评价,对比所选科普类文本语段在四个在线翻译平台的产出译文,得到的评分数据显示,各在线翻译平台译文质量从高到低依次为 DeepL、有道、搜狗、百度。在不同的科普类文本语段中不同的在线翻译工具译文质量不一,产出不稳定,DeepL 相对而言译文质量最稳定且质量较高,推荐作为翻译科普类文本的首选平台。但受在线翻译技术所限,在阅读时,读者仍需注意各种在线翻译工具翻译科普类文本可能出现的问题,建议提前了解术语、综合利用各种工具等方式辅助阅读。本文创新地融入英汉对比视角制定了科普类文本质量评价标准,以期能为科普类文本在线翻译译文质量研究提供一种新的评价思路。研究通过选取数学分析方法来提高标准权重客观性,尽可能地降低译文质量评价标准制定过程中的主观因素,但受评估模型、评价者水平、所选文本量所限,研究仍有待进一步深入。

参考文献:

- [1] 李星, 李伟山. 在线双语句库出现的误译——以铁路基础设施术语翻译为例 [J]. 山西能源学院学报, 2022, 35 (03): 100-102.
- [2] 陈胜, 田传茂. 在线翻译平台汉英翻译的问题及译后编辑——以石油地质文献为例 [J]. 中国科技翻译, 2021, 34 (01): 31-34+49.