

逻辑语义学视角下的连接词“and”的 句法错译分析

——基于汽车手册机器翻译的译后编辑研究

李阳春

(同济大学 外国语学院, 上海 200092)

摘要:通过对同济大学汽车学院的“汽车技术文献翻译语料库”中某一知名品牌汽车操作手册中随机抽取的2 500条机器翻译和人工翻译语料进行对比分析,并对其中420条含有连接词“and”的语料进行进一步探究,归纳了“and”翻译常见的错译类型,揭示了其背后所包含的逻辑语义关系,以期对机器翻译的译后编辑系统的升级提供一定的启示。经分析发现,连接词“and”的翻译,不仅要以句法为基础,而且要和语义结合起来,这样才能生成意义正确的翻译文本。

关键词:逻辑语义学;机器翻译;“and”翻译;译后编辑

中图分类号: H 059 **文献标志码:** A **文章编号:** 1009-895X(2019)04-0351-05

DOI: 10.13256/j.cnki.jusst.sse.2019.04.009

Syntactic Analysis of Mistranslation of the English Connective “and” from the Perspective of Logical Semantics

—A Study Based on Post-editing of the Machine Translation of Auto Manual

LI Yangchun

(School of Foreign Languages, Tongji University, Shanghai 200092, China)

Abstract: In this paper, a comparative analysis is conducted between machine translation and human translation in the 2 500 linguistic data which are randomly selected from a well-known auto brand's manual in Automobile Document Translation Corpus of Automotive College, Tongji University. In addition, 420 sentences including the English connective “and” are further analyzed. Common mistranslation types are summarized and logical semantic relations are discussed in the paper, hoping to provide some insights on upgrade of post-editing system for machine translation. It is found that accurate translation of connective “and” not only need to consider syntax, but also semantics. Only in this way can machine generate more accurate translations.

Keywords: Logical Semantics; machine translation; translation “and” into Chinese; post-editing

随着经济的快速发展和产品发布周期的日益缩短,对翻译交付的时间和速度提出了更高的要求。

以此为动力,机器翻译越来越成为高效翻译的重要组成部分,但是机器翻译固有的局限性影响了翻译

收稿日期: 2017-12-27

作者简介: 李阳春(1993-),女,硕士研究生。研究方向:应用语言学、外语教学教法。E-mail: Aprilian@tongji.edu.cn

的质量,阻碍了其进一步发展。提高机器翻译的准确性主要有两种途径,一是进一步完善机器翻译软件来提高所生成的正文质量,二是对机器译文进行译后编辑处理^[1]。完善翻译软件是最理想的解决途径,但是由于人类语言内部结构的复杂性,重新建构翻译规则不仅工程浩大,而且往往会触发新问题。因此,对机器译文进行译后编辑处理越来越引起人们的重视。译后编辑既可以由人工完成,一般称为 reviser,也可以利用软件来实现,一般称为 post-editor^[2]。本研究不聚焦于机器翻译的系统本身,而是关注系统产生译文后,如何建立一个独立的子系统进行译后编辑,从而提高最终译文的质量^[1]。

始于20世纪70年代初的逻辑语义学,又被称为形式语义学,是逻辑学与语言学的交叉研究学科。逻辑语义学以自然语言的意义为研究对象,利用现代逻辑为其研究工具,探索复合表达式的意义如何由其组成部分的意义组合而来^[3]。它与哲学、逻辑学、语言学、数学、计算机科学、人工智能和模型论等诸多学科息息相关,是一个横跨多个研究领域的知识体系^[4]。由于机器翻译规则的统一性,同一类型的翻译错误会反复出现,将逻辑语义学中的命题逻辑引入到机器翻译的译后编辑系统中,不仅能够有效理清句法之间复杂的逻辑关系,而且能够帮助改进译后编辑的算法,从而改进那些重复率高的翻译错误,提升机器翻译的质量。

一、研究设计

(一) 研究内容

本文通过对同济大学汽车学院的“汽车技术文献翻译语料库”中某一知名品牌汽车操作手册中随机抽取的英语原语、机器译文和人工译文各2500条语料进行对比分析,发现连接词“and”出现频率较高(占16.8%),但其翻译效果却不如人意。为此,本研究将在科技文本翻译的语境下,对连接词“and”的机器翻译错误进行描述和分析,归纳出机器在处理这类文本时所表现出的错误特征,并运用逻辑语义表达式为译后编辑系统的改进提供形式化的句法规则。

(二) 研究方法

本研究主要采用定性分析法来对随机抽取的2500条汽车手册中含有连接词“and”的语料进行

系统分析,并将机器译本和人工译本加以比较,从而对不同句法结构中连接词“and”的错译现象进行归纳总结。

二、不同句法结构中连接词“and”错译现象分析

在英语中,“and”是一个十分普通的连词,用来连接词、短语或句子,其基本意义相当于汉语的“和”或“并且”。这个词看起来很简单,有时甚至可以忽略或不译,但是在特定专业范围内,如本文涉及的汽车技术文本中,“and”有多种含义和用法变化,其翻译往往牵涉到整个句子的句法结构,能否对其进行准确恰当的分析直接影响到翻译效果和受众的理解程度。本研究希望通过分析归纳不同句法结构中“and”的错译现象,还原句子本身的逻辑语义关系,进而为译后编辑系统的改进提供一定的启示。

(一) “动词短语+多个并列宾语”的错译

例1 Source Language (以下简称SL): The automatic headlight beam level control system mainly consists of the AFS ECU, rear height control sensor, and two headlight leveling motors.

Machine Translation (以下简称MT): 自动前灯梁水平控制系统主要由AFS ECU组成,后高度控制传感器并且两盏前灯使电动机成水平。

Human Translation (以下简称HT): 前大灯光束高度自动控制系统主要包括AFS ECU、后高度控制传感器和两个前大灯光束高度调整电动机。

例2 SL: Remove the 3 clips and trim.

MT: 除去3个夹子并且整理。

HT: 拆下3个卡子和装饰板。

在例1中,动词短语“consists of”后接三个并列的名词短语“the AFS ECU”“rear height control sensor”和“two headlight leveling motors”,但机器未能识别动词短语后的三个并列宾语,而将名词短语“two headlight leveling motors”识别为与动词短语“consists of”并列的谓语成分,因而将“and”误译为“并且”。同样,在例2中,机器将“trim”识别为动词,从而造成句意不通。虽然“和”与“并且”都表示并列,但连词“并且”在汉语中表

示其连接的两个分句后者比前者在意义上更进一步^[5]。因此,要解决这一类型的句法翻译错误,笔者建议在译后编辑系统中将一些汽车手册中复现率较高的术语从词性上加以限制,从而使机器能够界定并识别出并列对象。

(二) 多个“动词短语+状语从句”的错译

例3 SL: Turns on when door is open *and* off when door is closed.

MT: 打开门营业什么时候和远,途经关闭什么时候。

HT: 车门打开时开关打开;车门关闭时开关关闭。

例4 SL: This switch turns off when door is locked *and* on when door is unlocked.

MT: 这开关关闭,到那时门锁住和在上到那时门被开锁。

HT: 车门锁止时此开关关闭;车门解锁时此开关打开。

例5 SL: This switch turns off when system is set *and* on when system is unset.

MT: 当系统被确定并且在系统是未装配的时导通时,这个开关闭。

HT: 系统设置时此开关关闭,系统未设置时此开关打开。

由以上三个例子可以看出,机器翻译在处理稍微复杂的句型结构时,其准确度大大下降,句法词序错乱。进一步分析便会发现,由于“and”之后省略了相同的动词“turn”,导致机器误以为“and”前后并列了两个形容词。在这种情况下,译后编辑系统需要使动词“turn”对副词“off”进行约束,从而使“on”和“off”搭建起联系,这样就很容易发现,“and”在这里表示对照,可译为“而”或直接用分号表示。

(三) 多个“并列小句”的错译

例6 SL: Engine switch off, all doors closed *and* driver side door key cylinder neutral position.

MT: 发动机关掉,全部门关上和司机门钥匙汽缸空挡的位置。

HT: 发动机开关关闭,所有车门关闭且驾驶员侧车门锁芯处于中间位置。

例7 SL: Engine switch on, luggage compartment door closed *and* luggage opener switch not pushed.

MT: 发动机打开,行李舱架门关闭和行李打开开关接通不按。

HT: 发动机开关置于ON位置,行李箱门关闭且未按下行李箱开启器开关。

经分析发现,机器将上述两个例子中的“and”翻译为“和”,主要是因为上述例子包含三个小句,这三个小句看似是并列的,其实不然。在例5中,“Engine switch off”实际上暗含一种条件,而例6中“all doors closed”和“driver side door key cylinder neutral position”是“Engine switch off”所引发的结果,因此,“and”实际上并列的只有后两个小句。可以进一步假设,如果这句话以完整的形式出现,即“(If) Engine switch (is) off, all doors (are) closed *and* driver side door key cylinder (is in) neutral position”,那么这句话中的内在逻辑关系便可理清,机器翻译的准确度也将会大大提高。例6属于同一类型的问题,所以在进行这类句型的翻译时,找出内在的逻辑关系就显得尤为重要。

(四) “并列句”的错译

例8 SL: If door is not completely closed *and* transmitter LOCK switch is pressed, this function makes buzzer sound for 10 seconds.

MT: 如果门没被完全关闭并且发送器锁开关被压,这个功能使蜂鸣器为声音10秒。

HT: 如果车门没有完全关闭而按下发射器锁止开关,此功能使蜂鸣器鸣响10秒。

例9 SL: When the rear door sunshade is operating *and* the engine switch is turned from on to off, check that the operation continues.

MT: 当后门阳伞正操作和从到离开上发动机开关被改变时,操作继续的检查。

HT: 后门遮阳帘工作且发动机开关从ON转至OFF位置时,检查并确认其继续工作。

在例8中,“and”的错译主要是由语义范畴造成的。在英语里,“and”常可将两种不同性质的陈述连接起来,表示对比或转折的意思^[6],而在汉语里,“并且”只能表示递进,不能表示转折。在以上例句中,“door is not completely closed”与“transmitter LOCK switch is pressed”是转折关系,此时的“and”可以由“but”替代。由于在英语里“and”有转折的作用,在本句里就可以用它将存在转折关系的前后两部分连接起来,而在汉译

时,译文里必须将“并且”改成“而”,这样句子前后两部分的关系才能被正确反映出来。这一例子给我们的启示是,要打破在机器翻译中原语词汇和目的语词汇的一一映射,将句法和语义相结合,才能提高翻译质量。而在例9中,“and”表示的却是递进关系,“the rear door sunshade is operating”和“the engine switch is turned from on to off”两个条件共同满足时,才需要“check that the operation continues”。由此可见,例8和例9虽然句型相似,但“and”表示的含义却大不相同,此时需要将“and”放在具体的语境下进行语义辨析。

(五) 复杂句中“并列多个主句”的错译

例10 SL: When the electrical key transmitter is brought near the electrical key antenna, the possibility of wave interference decreases, *and* it can be determined if wave interference is causing the problem symptom.

MT: 当电关键发送器被在电关键天线附近带来时,波干扰的可能性减少,并且如果波干扰正引起问题症状,它可能被确定。

HT: 将电子钥匙发射器带到电子钥匙天线附近时,波干扰的可能性减小,可以确定波干扰是否是导致故障症状的原因。

例11 SL: If performed on a speed tester, the front wheels do not rotate *and* the vehicle will determine that the vehicle is stopped.

MT: 如果在一速度检测器上执行,前轮不旋转和车辆将确定车辆停止。

HT: 如果用速度检测仪检测,前轮不转,则判定车辆停止。

例10和例11中,“and”看似连接了两个并列的主句,但是稍加分析就可以得出,这两个主句存在着因果关系。例10中,只有“波干扰的可能性减小”,才能“确定波干扰是否是导致故障症状的原因”;而在例11中,“前轮不转”才可以“判定车辆停止”。因此,“and”在这里的作用相当于“so”。在翻译时,应将“and”前后两部分之间的这种关系表达出来。在汉语里,“从而”这个词是表示上文是方法或原因,下文是目的或结果,用在此处来表达两个主句的关系十分恰当。

(六) “介词+多个并列宾语”的错译

例12 SL: For reassembly purposes, mark the

installation positions of the differential case bearing outer race *and* differential side gear shaft snap ring.

MT: 对重新聚集目的来说,标明承担外圈安装位置 and 不同边用具轴开口环的不同案件。

HT: 出于重新装配的需要,标出差速器壳轴承外座圈和差速器半轴齿轮轴卡环的安装位置。

例13 SL: Perform the tooth contact pattern inspection for the differential ring gear *and* drive pinion.

MT: 为不同戒指齿轮进行牙齿接触图案检查并且驱动翼。

HT: 检查差速器齿圈和主动小齿轮的齿接触斑点。

例14 SL: Support the rear suspension member with an engine lift *and* attachments.

MT: 支持后发动机的悬挂成员振奋和附件。

HT: 用发动机举升机和附加支撑块支撑后悬架横梁。

从以上三个例句分析可以得出,在“介词+NP+and+NP”这种结构中,机器往往不能识别出介词所管辖的范围,所以导致输出的翻译结果有误。在上述三个例子中,无论是介词“of”,“for”还是“with”,其约束范围均是由“and”并列的两个名词短语,而机器却误将第一个名词短语作为介词的约束对象,而将第二个名词短语机械地并列在后面。因此在译后编辑过程中,可以通过用逻辑表达式,将主干成分提取出来,再将这些成分进行排序,通过规则的运用限定这些成分的翻译次序,这样翻译的准确性将大大提高。

(七) 连接词“and”常见错译类型总结

通过以上的分析和归纳,笔者发现连接词“and”在汽车手册这一文本语境下的错译类型主要体现在三个方面:“and”连接多个名词性短语、“and”连接多个小句以及“and”连接多个句子,详见表1。

三、连接词“and”的命题逻辑表达式与译后编辑

逻辑语义学强调计算的思想、逻辑规则的采用和逻辑推演的套路,这使其成为分析自然语言句法语义构造组合的有力工具^[7]。由于科技文本的翻译需要精确的语义解释,而且其语用功能较弱,因

此, 笔者大胆尝试将数学和逻辑语言引入到机器翻译的译后编辑中。蒙太格认为, 每条语法规则都有一条与之对应的语义规则, 句法成分与语义成分之间是同态映射关系, 所有表达式的意义都是由其直接构成成分的意义组合而成, 语法规则组合一次, 对应的语义规则也组合一次^[4]。基于以上“and”错译现象分析, 可以发现句子结构一旦复杂, 机器

翻译的准确度就大打折扣, 然而, 在句法上要为每一句特例编制形式化的语法可能还不现实, 语言的形式化本身也是一项艰深的难题^[8]。因此, 笔者通过归纳连接词“and”的典型错译类型, 利用以下的逻辑表达式(见表2), 将这些句子隐含的逻辑关系显性化, 从而提高机器对汽车手册中含有连接词“and”语料的句法识别度, 改进翻译质量。

表1 连接词“and”常见错译类型总结

Tab. 1 Summary of common mistranslation types of connective word “and”

错译类型	举例	错译的具体体现
“and”连接多个名词性短语	例1~例2	动词或动词短语之后并列多个宾语, 机器翻译结果往往将最后一个宾语与动词或动词短语并列;
	例12~例14	介词后并列多个宾语, 机器不能识别出该介词所管辖的范围, 翻译结果常常只对毗邻的第一个宾语加以约束;
“and”连接多个小句	例6~例7	由于小句中缺少谓语动词, 其意义表达并不能通过形式完全呈现出来, 因此机器在处理这类句子时识别不出其内在的逻辑关系;
“and”连接多个句子	例3~例5	“and”前后连接了两个状语从句, 由于前后谓语动词相同, 所以原句往往会将第二个谓语动词省略, 这就使得机器在识别并列成分时出现问题;
	例8~例9	“and”前后连接了两个句子, 机器常常将这两个句子视为并列句, 因此译为“和”, 但是在具体语境下, 此处的“and”需要根据句意来进行相应的翻译;
	例10~例11	在复杂句中“and”连接了两个主句, 但这两个主句并非是简单的并列关系, 可能会存在因果关系, 但机器识别不出此种关系。

表2 “and”常见错译类型的逻辑表达式

Tab. 2 Logical expressions of semantic relations toward connective word “and”

“and”错误类型	举例	句型结构	逻辑表达式	符号说明
连接多个宾语	例1~例2	$V/VP + NP_1 + NP_2, \dots + \text{and} + NP_n$	$V/VP[NP_1, NP_2, \dots \& NP_n]$	V = 动词; VP = 动词短语; NP = 名词短语; P = 介词; “&” = “and”; C = 小句; S = 句子; “ $\rightarrow$ ”左边为条件, 右边为结果; “ $[\]$ ”内的内容优先翻译。
	例12~例14	$P + NP_1 + NP_2 \dots + \text{and} + NP_n$	$P[NP_1, NP_2, \dots \& NP_n]$	
连接多个小句	例6~例7	$C_1 + C_2 + \dots + \text{and} + C_n$	$C_1 \rightarrow C_2 \dots \& C_n$	
连接多个句子	例3~例5	$VP + S_1 + \text{and} + (VP) + S_2$	$[VP + S_1] + \& + [VP + S_2]$	
	例8~例9	$S_1 + \text{and} + S_2 + S_3$	$S_1 \& S_2 \rightarrow S_3$	
	例10~例11	$S_1 + S_2 + \text{and} + S_3$	$[S_1 \rightarrow S_2] \rightarrow \& S_3$	

由表2可以看出, 对连接词“and”常见的错译类型进行归纳后, 笔者用形式化的语言分别将每种错译类型进行描述, 并推导出其正确翻译的逻辑表达式。以表2中例10和例11来分析, “and”连接多个句子, S_1 和 S_2 作为共同条件可以导致 S_3 的结果, 而 S_1 又是 S_2 的前提条件, 便可推导出其正确翻译的逻辑表达式为 $[S_1 \rightarrow S_2] \rightarrow \& S_3$, 而这里的“and”便可译为“那么, 则”。

将这些逻辑表达式转换成机器翻译算法纳入到

译后编辑系统中, 将有助于提高机器翻译的准确性。

在此需要特别强调的是, 这些汽车语料中的“and”常见错译句法结构可以用逻辑表达式呈现出来, 根据这些表达式可以改进译后编辑的算法, 这虽然可以从句法结构的整体上提高机器翻译的准确性, 但在局部语境中, “and”究竟是翻译为“和”“并且”还是表示转折的“而”等, 还需人工进一步核对校验。

(下转第371页)

指导他们健康成长。二要立足少数民族大学生成长成才基本需求,努力为少数民族大学生加入到各种“师生共同体”中提供服务。通过指导少数民族大学生参加社团或开展创新创业实践、参与各种行业竞赛等,提升其综合素质,为他们成长成才提供指导。依托“师生共同体”采取定期召开座谈会、茶话会、谈心谈话以及书面交流等形式,掌握其最感兴趣的内容,及时了解学生的思想动态,使其树立科学、正确的世界观、人生观、价值观、国家观、民族观和历史观。

四、结束语

将少数民族大学生培养成德智体美劳全面发展的社会主义建设者和接班人是高校的根本任务。基于“师生共同体”视角开展少数民族大学生协同育人机制研究是推动少数民族大学生成长成才工作的重要举措,因此,还需要进一步加强“师生共同体”内涵、运行机制及效果研究。

参考文献:

- [1] 习近平.把思想政治工作贯穿教育教学全过程,开创我国高等教育事业发展新局面[N].人民日报,2016-12-

09(001).

- [2] 马克思,恩格斯.共产党宣言[M].北京:人民出版社,2014.
- [3] 马克思,恩格斯.马克思恩格斯选集(第1卷)[M].2版.北京:人民出版社,1995.
- [4] 中共中央文献研究室.建国以来重要文献选编(第5册)[M].北京:中央文献出版社,1993.
- [5] 毛泽东.毛泽东文集(第6卷)[M].北京:人民出版社,1999.
- [6] 习近平.习近平在北京大学考察时强调:青年要自觉践行社会主义核心价值观与祖国和人民同行努力创造精彩人生[N].光明日报,2014-05-05(001).
- [7] 卢梭.社会契约论[M].李平沅,译.北京:商务印书馆,2016.
- [8] 马克思,恩格斯.马克思恩格斯全集(第46卷)[M].北京:人民出版社,1979.
- [9] 马克思,恩格斯.马克思恩格斯全集(第23卷)[M].北京:人民出版社,1979.
- [10] 卡尔·雅斯贝尔斯.大学之理念[M].邱立波,译.上海:上海人民出版社,2007.
- [11] 龚放.大学“师生共同体”:概念辨析与现实重构[J].中国高教研究,2016(12):6-10.
- [12] 邓小平.邓小平文选(第2卷)[M].2版.北京:人民出版社,1994.

(编辑:程爱婕)

(上接第355页)

四、结束语

and虽然是一个小词,却是一个使用频率很高的词,其背后包含着复杂的句法逻辑关系,其语义远非我们所认为的“和”“而且”那样简单。在不同的语篇中,它有多种多样的词义,其具体译法也要根据不同的语境而定。在汽车技术文件这一限定领域内,能否准确翻译“and”,直接关系到译文的语气是否顺畅,逻辑是否严密,句意是否连贯,结构是否清晰。本文的上述诸例并未展现出and的全部语义功能,只是对连接词“and”的典型错译进行了深度的句法分析,初步尝试从逻辑语义的角度改进机器翻译译后编辑系统。但通过对“and”在汽车手册中常见的翻译错误类型进行归纳和分析,据此推导出正确翻译的逻辑表达式,缩小了对语言进行形式化处理的范围,有助于提高机器翻译的懂度和忠实度,并为机器翻译在这一领域的译后编辑软件的改良提供启示。

参考文献:

- [1] 李梅,朱锡明.译后编辑自动化的英汉机器翻译新探索[J].中国翻译,2013(4):83-87.
- [2] 李梅,朱锡明.英汉机译错误分类及数据统计分析[J].上海理工大学学报(社会科学版),2013,35(3):201-207.
- [3] 满海霞.关于逻辑语义学的句法系统[J].安徽大学学报(哲学社会科学版),2014,38(4):23-28.
- [4] 付艺,高芸.形式语义学发展概述[J].江西社会科学,2013(11):20-24.
- [5] 邵启祥.谈科技英语中and的翻译[J].中国科技翻译,1995,8(1):18-22.
- [6] 徐金榜.小字并不小,译时需推敲——谈英语连接词and的翻译[J].解放军外国语学院学报,1996(3):47-52.
- [7] 邹崇理.逻辑语义学的方法论创新[J].湖南科技大学学报(社会科学版),2016,19(6):27-34.
- [8] 罗季美.机器翻译句法错误分析[J].同济大学学报(社会科学版),2014,25(1):111-118,124.

(编辑:朱渭波)