



Minerva Access is the Institutional Repository of The University of Melbourne

Author/s:

Pym, A;Fan, M

Title:

Aspects of machine translation. An interview with Professor Anthony Pym

Date:

2021

Citation:

Pym, A. & Fan, M. (2021). Aspects of machine translation. An interview with Professor Anthony Pym. *Shandong Foreign Language Teaching*, 42 (3), pp.3-12

Persistent Link:

<https://hdl.handle.net/11343/280526>

DOI: 10.16482/j.sdwy37-1026.2021-03-001

# 机器翻译面面观

## ——皮姆教授访谈录

范梦栩<sup>1</sup> 安东尼·皮姆<sup>2</sup>

(1. 中央民族大学 外国语学院, 北京 100081

2. 墨尔本大学 语文及语言学学院, 墨尔本 3010)

**摘要** 近年来,信息技术日新月异,人工智能发展迅猛,翻译技术成为翻译研究领域的热门话题。知名翻译学者安东尼·皮姆(Anthony Pym)教授多年来一直从事与翻译技术相关的研究,其教学实践同样重视与机器翻译相结合。本人受《山东外语教学》之托,邀请皮姆教授就翻译与技术的关系、机器翻译对职业翻译的影响、机器翻译的质量评价和伦理、机器翻译对译者培训和语言保护的积极作用等问题深入阐述了他的观点。希望此次访谈能够破除某些对机器翻译的成见,并对其未来的应用和研究带来一定的启发。

**关键词** 机器翻译; 翻译技术; 翻译教学; 译者培训; 安东尼·皮姆

**中图分类号** H059 **文献标识码** A **文献编号** 1002-2643(2021)03-0003-10

### Aspects of Machine Translation: An Interview with Professor Anthony Pym

FAN Mengxu<sup>1</sup>, Anthony Pym<sup>2</sup>

(1. School of Foreign Studies, Minzu University of China, Beijing 100081, China

2. School of Languages and Linguistics, The University of Melbourne, Melbourne 3010, Australia)

**Abstract:** The rapid development of information technology, especially artificial intelligence, has triggered breakthroughs in translation technology, currently a hot topic in translation studies. Professor Anthony Pym, an internationally-renowned figure for his study of translation, has been researching this topic and applying machine translation to his teaching. Entrusted by *Shandong Foreign Language Teaching*, we invited Professor Pym to expound his views in depth on the relation between translation and technology, the impact of machine translation on professional translation, the quality evaluation and

收稿日期: 2021-03-09; 修改稿, 2021-04-07; 本刊修订, 2021-05-06

基金项目: 本文为国家社科基金后期资助项目“中国电影走出去: 一种海外研究的视野”(项目编号: 17FYS021)的阶段性成果。

作者简介: 范梦栩, 博士研究生。研究方向: 翻译史、翻译与文化。电子邮箱: fmx0515@163.com。

安东尼·皮姆, 博士, 教授。研究方向: 翻译理论、翻译伦理、翻译史、机器翻译。电子邮箱: anthony.pym@unimelb.edu.au。

ethics of machine translation, and the role it plays in translator training and language protection. It is hoped that this interview can help eliminate some misunderstandings around machine translation, and provides insights for its future application and research.

**Key words:** machine translation; translation technology; translation teaching, translator training; Anthony Pym

范梦栩(以下简称范):皮姆教授您好,非常感谢您接受此次采访。您的研究领域非常广泛,对翻译技术等相关问题的挖掘更是翻译研究的前沿之一。科技影响了我们生活的方方面面,也深刻影响着翻译实践和翻译研究。现代翻译研究越来越依赖科技,有的学者认为这种取向符合数字人文趋势,值得推广;也有学者对此提出质疑,认为这种取向不够重视对翻译本体、也就是对文本的研究。对于翻译研究的“科技取向”您怎么看?

皮姆教授(以下简称皮):这个问题很好。在人文学科中,我们应该对所有可用的技术持开放态度,尤其是将大数据方法纳入研究。其他学科正在这样做,翻译研究为什么不呢?利用这些技术,我们的部分成果可以呈现为好看的图表。但是如果研究的问题和假设都没有什么意义,再漂亮的图表也是无用的。因此,我认为仅仅使用技术是不够的,必须使用技术来解决问题——当然前提是要有有趣的研究问题以及巧妙的解决方法。另外必须承认,我认为翻译研究的对象不一定是文本。长期以来,我都是“翻译研究以人为本”的推崇者之一,尤其最近我对研究同样是人的翻译读者很感兴趣。数字人文的陷阱是,我们往往让数字控制了自己的思维。我们不应忘记,在人文学科中,我们是与人的特质、人与人之间的交流、人的思想以及定性的数据打交道。眼动追踪相关的研究就是一个很好的例子。这类研究产生的数据通常比回答研究假设多得多,但其中也存在数据丢失、技术和方法论等方面的巨大问题,因为有时候眼睛所至之处并不代表大脑所想,即眼与脑并非时刻一致。尽管如此,图表看起来还是很棒的,我很乐意使用图表。

范:机器翻译仅有七十年左右的历史。近二、三十年来,随着神经机器翻译的发展,机器翻译质量显著提升。华为公司翻译中心提供的数据显示,2018年该中心日均纯机器处理、无人工译后编辑的翻译交付量已达1亿字符(任文,2019),一时间机器翻译会取代译员的说法甚嚣尘上。有些人认为这是杞人忧天,有些人则嗅出了危机感。您认为机器翻译的发展对当前和未来的翻译从业者意味着什么?

皮:你刚刚提到的华为的数据我很感兴趣。他们是如何知道机器翻译结果没有经过人工译后编辑的呢?有什么方法可以得知吗?翻译后的文本有人去读吗?这样庞大的字符翻译量中包含多少有效信息呢?例如,我可以用机器翻译将整个网页译为中文,这是很多字符,但是我也许只需要其中的三个字符组成的一条信息。华为提供的2018年的数据甚为惊人,实际上谷歌翻译也发布过类似的数据。但是我认为我们对待这些数字时一定要格外谨慎,要思考它们究竟代表了什么。首先,这些数字真的能反映机器翻译的使用情况吗?因为如果可以免费把字符交给机器翻译系统处理,那么人们使用时就不会感到任

何交付数量上的限制。第二个问题,也是最重要的问题是,如果机器翻译不是免费的,那么人们会把这些字符交由专业笔译员或口译员处理吗?例如,2000年的一项调查显示,人们很少请职业译员翻译在日常生活中能用机器翻译处理的字符。因此仅就数字而言,我认为机器翻译并不会对专业口笔译人员构成直接威胁。现在的情况是,整个翻译需求的蛋糕变得更大了。尽管译员所占的份额少于所有翻译字符数的1%,他们的工作量依然在不断增长。从我们掌握的统计数据看来,在美国、加拿大、德国等地,职业口、笔译员的雇用数量都在增长。也有来自中国的数据,称职业译员的数量以每年10%的速度增长。因此我尚未见到任何实质性的证据可以证明机器翻译正在取代职业翻译。人工翻译与机器翻译的竞争,本质上是人工翻译与机器翻译加译后编辑、或者说与机器翻译加人工翻译之间的竞争。一旦翻译人员习惯了机器翻译,他们的工作效率以及术语和固定短语的翻译质量将大大提高,我认为这一点不难证明。因此,我认为问题的重点不是人类与机器的对立,而是人类如何有效地使用机器。

范:有人说,机器翻译在科技翻译(非文学翻译)领域或大有可为,但是利用机器进行文学翻译则是“不可能完成的任务”。请问对此您怎么看?

皮:我认为最好的办法是试试看,也就是运用实证的方法来探究我们不知道的问题,看会发生什么。我曾在课上让学生们用翻译机器将沃尔特·惠特曼(Walt Whitman)的英文诗译成中文,最终我们发现许多机器翻译结果都相当成功,反而是部分人工翻译结果不尽如人意。可见对于某种特定程度的复杂性和难度而言,完全没有理由说机器翻译的结果比人工翻译的差,也没有任何理由不让译员使用翻译机器。另外,我也不认为文学与其他类型的文本有什么本质上的区别。惠特曼就是一个例子,他的语言非常简单。自19世纪末以来,大多数西方文学都有语言越来越生活化、口头化的发展趋势。我不太了解中国文学的情况,但我推测可能也是如此。以前我们称之为经典的那些文学作品往往与日常生活相割裂,如今文学与日常生活越来越相近了,尤其是诗歌。因此毫无疑问,文学对机器翻译并无特殊。不过还有一点很有趣。我在墨尔本的课堂上、我的同事们在西班牙的课堂上都进行过相关试验,我们发现当学生使用机器翻译系统来辅助文学文本的翻译时(即先用机器翻译然后再进行译后编辑),他们的翻译质量普遍提高了。通过计算基本错误——翻译错误和语言错误等的数量,可以证明使用机器翻译并进行译后编辑的学生比不使用的学生表现更好,也译得更快一点。但是,当我们问这些学生“你是否会继续使用译后编辑”时,很多人说“不,我不喜欢对文学文本进行译后编辑”或者“机器无法翻译文学作品”。也就是说,即使我们可以用数字告诉他们借助机器的翻译效果更好,在思想认识上他们仍然不愿使用机器翻译文学文本。可见问题不在于语言事实,而是在于人们在固有印象中认为文学是常人无法企及的语言的最高境界,于是对机器翻译持有抗拒心理。

范:也就是说这个问题与信任有关?译者不够信任机器翻译?

皮:没有人应该完全相信机器翻译,但是我们可以证明人们能用机器翻译达到目的。

这是一个可信度的问题,同人与人之间的关系是一样的。如果你自问“我能信任这个人吗”,同时很快发现这个人犯了一两个小错误,可能你就会告诉自己“他肯定还有很多其他错误,我不能信任这个人”。这就像复杂理论( complexity theory)中的蝴蝶效应,一两件小事就能将信任变成不信任,因为建立信任是很困难的,打破信任却很容易。这便解释了为何人们即使会采纳机器翻译的一部分结果,却仍然认为机器翻译不好用。对译者来说,即使他们相信源语文本和目标语文本中那些具体的东西,他们也不一定信任机器翻译本身。打个比方的话,也许可以这么想:我并不相信卖东西的人,但是如果价格还可以,物品也实用,我仍然会从他那里买东西。

范: 2019 年出版的《劳特里奇翻译与科技手册》( *The Routledge Handbook of Translation and Technology* ) 中,您撰写了“翻译质量”这一章( Pym, 2019)。您在文中提到,人们最常提出的问题就是“机器翻译能不能翻译得像人类一样好”,而往往翻译“好”与“不好”的标准是不确定的。那么在实际研究中,究竟应该如何评估机器翻译的质量呢?

皮: 首先我需要解释一下“不确定性”( indeterminacy)。我认为对于一个给定的翻译问题,并非只有一个合理的解决方案。由于语言不确定性的本质,语言间不存在机械的因果关系。这意味着针对同一个原文本,不同的译员会给出不同的翻译文本,也就是给问题找到不同的解决方案,而所有方案都是可接受的。我们这些一直与翻译打交道的人都很清楚不确定性的存在,即有很多不同的标准可以使用。而我们评估机器翻译的质量时,却总是使用人类的翻译作为参考答案,并认为它是唯一的标准。这就是问题所在。不确定性给质量评估带来的困难不单单是机器翻译的问题,也是翻译研究总体上的问题,更是语言评估中存在的难题。因为不确定性是语言的本质属性之一,而不是特定于翻译的。

我在研究中会尽可能回避翻译质量这个问题,因为实在是困难。就我而言,质量的含义是多重的。质量评估就像审美“情人眼里出西施”嘛。不同的人会在不同的地方、不同的事物中发现美,发现不同的美。那么具体该怎么做呢?一种方法是在承认存在不同标准的前提下,人为设定一个标准。我们可以根据自己的研究目的进行说明,选用某一种翻译作为正确答案,而其他的都是偏差值。此外,如果几位评估者能够在某些标准上达成共识,或同意某些译法是不足甚至错误的,我们也可以利用他们的意见辅助判断。另一种方法是观察读者和他们的实际反应,即认为翻译的质量取决于人们阅读后的感受。我们可以通过问卷调查了解读者对译文的理解或者探寻他们下一步的打算,例如在销售场景中,读者愿不愿意购买某个译本可以作为衡量其质量的标准。另外,基于之前提到的信任问题,我们也可以让人们阅读不同的翻译,然后问他们信任哪个、不信任哪个以及为什么。由此你也许会发现,一些小错误就会破坏译文整体的可信度,导致读者心目中译者及其翻译工作的骤然贬值。总之,我认为衡量译文质量的方法有很多,而我们选择的方法应该与自己要在研究中解决的问题相对应。

范: 同样在这一章中,您还提到了一个非常有趣的话题: 翻译技术、尤其是机器翻译对

译者工作幸福感和满足感的影响。就现有的研究而言,请问您认为是否存在一个平衡点,使译员在机器翻译的帮助下实现最大程度的满足?

皮: 这个问题很好,但是没有确切的答案。很多年前,我有一个来自香港的博士生廖凤明(Liu Fung Ming, Christy),她对译者的幸福感进行了研究。幸福感是可以衡量的,使公民感到幸福是国家的义务之一,因此对幸福感的研究很有意义,也已有大量前期成果。廖发现,译者的幸福感很大程度上取决于他们的个性。不同的人乐于做不同的事情,幸福并没有一定之规。其实这一点我们本该心知肚明,但不知为何我们在讨论中往往不承认它。廖的研究发现,有些译员很乐意与他人面谈,面对面解决问题,他们是真正的“调解人”(mediators);另外一些译员则喜欢藏在计算机后面解决技术问题,查找资料以获得特定术语的精确匹配项;还有很多译员介于两者之间。

那么机器翻译如何影响译者工作满足感呢?要看情况。例如,有些人比较擅长直接在电话上为人解决问题,那他们不会喜欢机器翻译,因为它会分散他们的注意力,成为障碍,使他们与实现交流的目的背道而驰。有些人在处理任务的过程中,能发现并更正机器翻译的错误,改进数据库并从中获得极大的满足感和幸福感。我虽然不是这样的人,但的确有人是。如今很多翻译毕业生都在进行清整数据库的工作,即检查并验证输入到机器翻译数据库中的信息。对我来说这实在是乏味得可怕。有一次我与莎伦·奥布莱恩(Sharon O'Brien)讨论这个问题,她可以算是翻译技术方面最顶尖的学者了。我说“莎伦,你看,很多翻译毕业生不得不去做这些无聊至极的苦力活儿。”她却说“我完全不觉得啊。”她在欧洲委员会的第一份工作就是有关数据库的。在那里她学习了技术原理,了解了机器翻译的工作方式以及改善人机交互的方法,所以她并不同意我的意见。她与我性格不同,研究语言的方式也不一样。数据库工作令她感到心满意足,并成为她进入翻译技术研究领域的切入点。这再次说明解决方案因人而异。在我以前的学生安娜·格伯罗夫(Ana Guerberof)的研究中,我们发现首次进行译后编辑的人通常会讨厌它,但一旦习惯之后,讨厌它的程度就会降低。就像我们第一次尝试任何事情一样,这只是一种针对新事物时最初的负面反应。

范: 在科技发展日新月异的背景下,职业译员的伦理问题引发学界关注。2020年《山东外语教学》也刊发过“翻译伦理再考察”专题,探讨如何建立译员职业道德考核评价机制(任文,2020;王巍巍、余恽,2020)等问题。请问您怎样认识机器翻译伦理?

皮: 这很难讲。总的来说,我们的社会中存在许多与大数据和自动化有关的重要问题,我认为就翻译而言这些问题并无不同。正如《中国语言服务行业规范》(2019)中所规定的,显然译员不能将机器翻译的结果作为人工翻译来提交。但这只是良好的商业惯例,即无论产品是什么,都要在交付产品时保持诚实。因此,我不确定是否存在与翻译相关的特殊问题。但是,我很乐意概括并讨论翻译界正在发生的事情。

对我来说,机器翻译伦理中最大的问题是数据库的构成,因为它不像是人们把DNA

或自己的照片放入一个巨大的数据库中,而是拿走了我们的劳动成果。当我们在线工作时,我们的翻译会保留在网站上,然后网络爬虫会提取这些翻译,放入数据库中;我们使用免费的在线翻译软件时,软件会将我们的操作记录收入数据库中。机器翻译系统中可供贸易的存量——其真正有价值的部分——取决于数据库的质量。机器翻译依靠的不是语言学而是数学,不仅需要数量庞大的数据库,更需要由人工生产高质量的配对。换言之,我们免费提供了劳动力。这是不是一个伦理问题?如果你是一位坚定的马克思主义者,就肯定不会否认这一点。就像资本家建了工厂,工厂利用了工人的劳动力,工厂归资本家所有,工人自愿为资本家提供了剩余价值,成了资本家获得资本的方式。数据库正在做的事与此别无二致。另一方面,我认为作为工人,即译者、语言工作者或者相关研究者,如果可以从这笔资本中受益,那也无可厚非。但是请注意,这其中也涉及数据管理、隐私以及为交换某些东西而无偿提供数据等伦理问题。这样的伦理问题充斥着整个社会,而翻译只是其中的一个领域。

我还想提出另一个重要的伦理问题,因为我认为这对中国尤为重要,而涉及中国的一切都是非常大的数字。技术使我们能够翻译一切,而技术界的理念是,如果能够做,那就去做。现在有了免费的在线机器翻译,我们也许更会这么想。然而,正如迈克尔·克罗宁(Michael Cronin)所指出的那样,从生态学的角度讲,无限制的机器翻译是不可持续的(Cronin, 2017)。无论是从翻译学还是生态学的角度看,我们不能持续不断地利用地球的资源来存储数据,也不能不断建立数据库使机器持续产生琐碎无用的结果。终有一个临界点会让我们不得不开始思考“不翻译”,或者说,思考哪些文本值得利用地球的资源进行翻译而哪些则不需要。我们没有马上意识到这一点,是因为世界上并不存在这样一个实体的“计算机农场”,肉眼无法看到其中正在消耗能源接受处理的数据。但是这个生态问题是确实存在的,我们应该思考如何在真正需要技术的时候使用它,而不是仅仅因为它存在就去肆意使用。我自己关于伦理的书(Pym, 2012)中也提到了这一点。我提出了一个根本性问题:究竟为何要翻译?实际上,翻译是需要理由的。我在书中举例说明了为什么我认为有些文本不应该翻译,不是因为低俗淫秽或其他类似的理由,而是因为翻译它不会产生任何益处。既然不存在协作,也就没有必要交流。我认为我们需要进一步思考为什么要翻译以及为什么不翻译。

范: 刚刚您的回答引人深思。尤其是数据管理问题,很多时候我们并没有意识到自己的翻译成果被用在了其他地方。

皮: 是的,这种情况一直都在发生,甚至比技术产生得更早。有时候我会有一些数据库、术语库里看到自己的翻译,兜了一圈又回到我身边。只要这份劳动是用于交换,并且人们能清醒地意识到自己的成果被利用并同意说“我愿意把这个给你用于交换”,对我来说就没有问题。但是如果交换并不平等,那我就不能认可了——实际上这就等于我们被抢劫了。我再举一个例子。大概 8 年前,世界各地的翻译公司收集了大量语言数据,用有

效的双语语料建立了数据库。他们开发了一个“人类语言项目”(the Human Language Project), 试图将全部数据汇集起来, 齐心协力解决翻译问题。之前有个类似的项目叫“人类基因组计划”, 是将全世界的科学家聚集在一起, 收集他们的资源绘制人类基因组的图谱, 然后把它用于全世界的医学研究, 造福所有人。听起来很棒是不是? 但从2016年开始, 人们逐渐意识到对数据库来说最重要的是它有多干净而不是它有多大, 我们需要的是高质量的数据库。但这些翻译公司可不会为了全人类的利益而免费提供服务, 他们开始出售那些“属于他们”的东西: 开发人类语言项目的组织 TAUS(Translation Automation User Society) 开办了一个市场, 供人们买卖数据库或所谓的“语言资产”(language assets)。由于技术的进步, 这个曾经伟大的人道主义项目变成了市场, 一部分人的劳动成果被自以为拥有这些成果的另一部分人拿去出售了。这就是一个伦理问题。

范: 我有幸在墨尔本大学访学期间旁听过您的“翻译与技术”课程, 请问您是如何想到要将机器翻译融入翻译教学的? 对翻译专业的学生来说, 培养语言能力是重中之重, 那么翻译技术的教学应放在什么阶段、用怎样的方式进行最为合适? 去年您还发表了一篇有关译员培训的文章(Pym, 2020)。请问您认为机器翻译会给译员培训带来什么样的帮助呢?

皮: 这个问题和“为什么要翻译”差不多。不应该因为技术就在那里而使用技术, 这点对翻译教学和译员培训是一样的。如果学生已经在使用机器翻译了, 那么就必须要教他们如何使用。否则他们虽然用了但可能使用不当, 或者因为没有意识到机器翻译的局限性而过于信任它。我实施教学和培训时会从译后编辑入手, 先让学生观察哪些地方出了问题, 需要改正什么, 为什么要改正它; 然后进行逆向思考, 搞清楚为什么机器翻译会得出错误答案; 最后进行译前编辑, 看看是否可以改写源语文本让机器翻译得更准确。仅通过这些, 我们就学到了很多有关机器翻译的知识, 同时也可以教授译后编辑和译前编辑这两种技能。我这样做时, 不会告诉学生机器翻译是好是坏或者进行任何事前判断, 他们必须自己去发现。每年我都做相同的活动, 但是每年都能得到不同的结果。这其中的原因是机器翻译一直在进步, 翻译结果变得越来越好。教学应当鼓励学生自己去发现机器翻译的优点和缺点。

范: 您能简要地向我们介绍一下目前机器翻译研究领域的热点问题吗?

皮: 我不知道(笑)。几年前我写过一篇论文, 其中提到机器翻译研究的问题是机器翻译或者翻译记忆库两年内就会发生很大的变化, 但是发表学术论文就需要两年左右的时间, 研究本身可能也需要两年时间; 如果是博士论文则需要四或五年的时间, 那么从着手研究到将论文发表成书, 可能花费五、六年甚至七年。因此最好不要研究机器翻译, 因为机器翻译和翻译记忆库在四、五年后早已与今天不同。如果还是想研究的话, 必须抽象到一定高度。不要研究目前已知的东西, 比如网络在线翻译的质量、西方和中国的神经机器翻译系统的比较之类, 它们很快就会变化。我认为研究者需要退后一步, 把眼光放在一

些不受时间影响的主题上。例如人对机器翻译的抗拒情绪就很有趣:人们为什么抗拒?如何抗拒?这种抗拒会如何变化?查阅人们对机器翻译的评论,解释其背后的原因,将是带有社会学意义的实用话语分析。

还有一个非常有趣的课题值得学者研究,也能为行业提供帮助,就是那些缺乏数据的语言。机器翻译对于使用者众多的语言和有大量电子资源文本的语言相对有效,因为这些语言有大量电子资源支撑数据库。但是对于许多较小众的语言,资源稀缺,鲜有电子形式的记录,因而被排除在机器翻译本可以带来的便利之外。因此我认为,深入研究这些语言并思索如何建构数据库,如何使其变得可靠,如何在特定领域中使用非常小的数据库来产生良好的机器翻译结果,都会是有趣的研究。我来自澳大利亚,这里有一种叫努嘎语(Noongar)的土著语,是一门几乎灭绝的语言,会讲的大约不到 100 人,但是人们正在努力复兴它。不会讲努嘎语的人,如果他们的父母也不会的话,就会去找仅存的还会这门语言的人学习,并用努嘎语做各种各样的事,比如将莎士比亚翻译成努嘎语。正巧我今天还看到新闻说,因为他们都喜欢李小龙,就把电影《精武门》译成了努嘎语,所有去看功夫打斗的孩子们都会从电影中潜移默化地学习努嘎语。他们还开发了努嘎语的百科全书“Noo-ngarpedia”,这部百科的生成方式类似于维基百科(Wikipedia),但使用的是努嘎语。他们还在建设努嘎语的知识库,这需要努嘎语的数据库和术语库。这是我国土著语言的例子,世界上还有许多较小、濒临灭绝的语言都在进行这样的工作,而技术确实可以在这场斗争中提供帮助。因此,我认为不仅仅是机器翻译,翻译本身能对语言复兴(language revival)起到的作用是很有前景的研究领域。研究应当如何更好地复兴一门语言,是我和墨尔本大学的同事做过的最美妙的事情。我们做过一种特别神奇的工作,就是利用语言学信息使某些语言起死回生,或者教给孩子们连他们的父母都不会的已逝的语言。我是澳大利亚人,但祖先在欧洲。我们曾经入侵了这个大陆。出于入侵者的愧疚,我认为我们必须这么做,这是我们摆脱愧疚的一种方式。如果技术能协助我们,实在是求之不得。

范:请问您是否关注过中国学者在机器翻译研究领域的成果?您曾经提过“语言距离”(distance between languages)这一概念,中文的特性会使机器翻译研究更复杂吗?

皮:中文方面我可不是专家,所以也无法评价。我只知道中文也是一种主谓宾结构的语言,和英文相同,但与大多数亚洲语言不同。所以中文和英文互译时,句法方面往往不会像韩文和日文那样出现重大问题。另外,中文拥有庞大而且干净的数据库,因此我们使用汉英数据库(汉英机器翻译)时可以得到不错的结果。基于这两个因素,即句法的相似性和庞大整洁的数据库,汉英互译的结果不像人们想象的那么差。关于另一个问题,首先我想问你,你读过加拿大翻译学者的文章吗?

范:说来惭愧,读论文的时候,我好像很少关注学者本身是从哪里来的。

皮:那么法国学者的呢?或者是德国的、乌克兰的或澳大利亚的?你应该明白我的意思了。我认为翻译研究不是奥林匹克运动,也没有那么一支“中国队”。我读过的一些论

文署名像是中国人,但其作者可能来自世界的任何地方。墨尔本和澳大利亚的其他地区有我的中国同事在研究机器翻译及其应用。在机器翻译和翻译技术研究中心都柏林,也有很多中国学生做着出色的研究。香港是翻译研究重地,有很多翻译项目,数不清的中国学者都发表了成果。所以这个问题对我来说很有意思,因为我可能不会问关于法国人的、乌克兰人的或者加拿大人的研究。再比如,我与北京外国语大学的任文教授讨论过伦理和机器翻译的问题,但我并未觉得我是在进行一种跨越国界的交流,而只是以翻译学者的身份与她沟通。因为我们正在研究相同的问题,大多数情况下遵循着相同的学术传统,我们之间存在某种共通的文化。

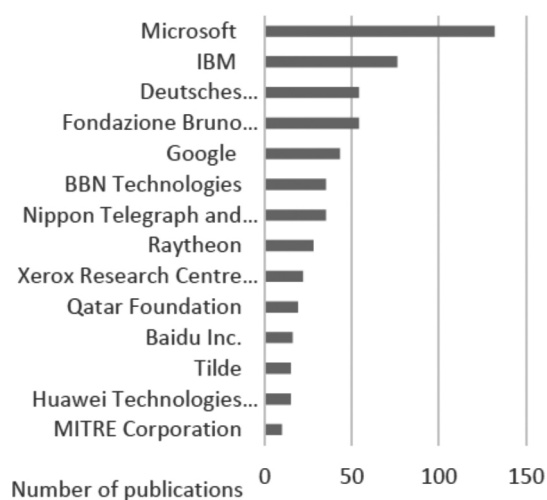


图1 Scopus 数据库中有关机器翻译的出版物数量及获资助情况  
(2010年-2018年6月)(Vasiljevs et al, 2019: 3)

今天我看到一份有关机器翻译的出版物的数据(图1),该图显示,微软和谷歌赞助出版了很多学术出版物,然而来自中国的公司排名并不高。但是所谓“来自中国的研究”究竟是什么?作者有中文名字就算吗?论文是用中文写的就算吗?还是为了中国公司而完成的成果就算在内?百度是中国公司中排名最靠前的,却仅排在倒数第四位。因此在这个领域,并非有人故意忽略中国学者的研究,只是恰好有更多来自西方的研究。上次布里斯班国际翻译协会(Fédération Internationale des Traducteurs, FIT)的会议上我做了一个主旨发言,在我之前,赞助大会的一家中国公司的代表就大数据技术以及中国在该领域如何领先世界作了重要演讲。这是国际翻译协会的世界级会议,可以证明世界并没有忽略中国的研究成果。我觉得有趣的是,该演讲中强调了大数据的重要性,而在我看来,最重要的是干净的数据与这些数据的所有权。因此,对于同一件事,我们的关注点可能不同,但无论对方来自哪里或用什么语言,我都很高兴与他们坐下来谈论翻译。从许多方面看,这并没有太大的区别。

范: 如果有学者想要从事机器翻译相关研究, 您可以在方法论方面提供一些建议吗?

皮: 其实就是刚刚提到的那一点: 不能把眼光放在那些变化速度太快的对象上, 而是应该抽象到一定高度, 抽身出来着眼大局。细节问题可以交给这个行业的专家, 或者真正关注具体数据以及其他动态的那些公司。对于学者来说, 当我们讨论译者抗拒机器翻译及其原因这类问题时, 与其他实证研究或思想认识领域的研究没有什么不同。因此我认为翻译技术的讨论在方法论领域并没有什么特殊之处。

范: 在访谈的最后, 请问您能否推荐一些有关机器翻译的必读经典和前沿之作, 供感兴趣的研究者们参考学习?

皮: 不能(笑)。正如之前提到的那样, 出版一本书要花很久, 机器翻译却瞬息万变。我认为这个领域完全可以直接通过实践体验, 因为使用翻译工具轻而易举。我们可以亲自体验机器翻译系统, 还可以查看运营机器翻译系统的公司发布的最新文章。当然, 也不要太把他们的宣传当真, 毕竟他们的目标读者群并非翻译研究者。

范: 非常感谢。记得您曾经在一次学术报告中说, “神经机器翻译对当今社会最大的贡献就是每个人都能使用它, 并且从中得到乐趣。机器翻译并不是为了获得唯一确定的信息, 而是一种探索世界的方式”。我非常赞同。您的观点与翻译活动本身的价值也有相通之处——扩展人们对世界的探索。希望也相信今后的科技能够不断帮助人们实现这个心愿。再次感谢您分享关于翻译技术和机器翻译的见解, 期待今后继续与您保持交流。

#### 参考文献

- [1] Cronin, M. *Eco-Translation: Translation and Ecology in the Age of the Anthropocene* [M]. London: Routledge, 2017.
- [2] Pym, A. *On Translator Ethics: Principles for Mediation Between Cultures* [M]. H. Walker (Trans). Amsterdam: John Benjamins, 2012.
- [3] Pym, A. Quality [A]. In M. O'Hagan (ed.). *The Routledge Handbook of Translation and Technology* [C]. London: Routledge, 2019. 437 - 452.
- [4] Pym, A. For a sociology of translator training [A]. In J. F. Zhao (ed.). *Translation Education: A Tribute to the Establishment of World Interpreter and Translator Training Association (WITTA)* [C]. Singapore: Springer Singapore, 2020. 9 - 19.
- [5] Vasiljevs, A. & I. Skadinaetal. Competitiveness analysis of the European machine translation market [A]. In *Proceedings of Machine Translation Summit XVII Volume 2: Translator, Project and User Tracks* [EB/OL]. 2019. <https://www.aclweb.org/anthology/W19-6700.pdf> [2021-02-08]
- [6] 任文. 机器翻译伦理的挑战与导向 [J]. 上海翻译, 2019, (5): 46 - 52 + 95.
- [7] 任文. 新时代语境下翻译伦理再思 [J]. 山东外语教学, 2020, (3): 12 - 22.
- [8] 王巍巍, 余恽. 译员职业道德考核评价机制探索 [J]. 山东外语教学, 2020, (3): 23 - 31.
- [9] 中国翻译协会. 译员职业道德准则与行为规范 [EB/OL]. 2019. [www.tac-online.org.cn/uploadfile/2020/1110/20201110103058580.pdf](http://www.tac-online.org.cn/uploadfile/2020/1110/20201110103058580.pdf) [2021-02-08]

(责任编辑: 杨彬)