

数据驱动学习模式下会展翻译教学探索

梁 爽

(重庆第二师范学院 外国语言文学学院, 重庆 400067)

摘要:数据驱动学习是一种基于语料库数据的外语学习方法,给会展翻译教学带来了新的突破。网络搜索和单语语料库与会展翻译教学相结合,是一个鲜有人探索的领域。在课堂上利用免费获取的英语单语语料库 COCA 和 BNC 以及规模庞大的 Wikipedia,能改变会展翻译的传统教学模式,为各种翻译课程的改革提供参考。

关键词:数据驱动学习;单语语料库;会展翻译教学

中图分类号:H319.3

文献标识码:A

文章编号:1008-6390(2017)04-0113-05

一、引言

按照《翻译专业本科教学质量标准》,翻译专业旨在培养能够胜任多领域工作任务的通用型翻译专业人才。从翻译市场需求的实际情况来看,翻译从业人员大多从事商务等非文学领域的翻译工作。因此,为了适应新形势的需要,翻译专业教学应该转向“英语翻译+专业知识”的高素质通用型人才培养模式。

中国政府大力推进“走出去”“引进来”的战略方针,会展行业也迎来了蓬勃发展。“会展活动是人类社会、经济、文化交流发展到一定阶段的产物。”^[1]中国已成为全球发展最快的新兴会展市场之一。会展经济要进一步发展,翻译的作用就绝不容忽视。因此“会展翻译”课程的开设符合翻译专业的培养目标,迎合了新形势下培养通用型翻译人才的需求。

虽然目前有不少高校的本科翻译专业也开设了会展翻译类板块课程,但是这些课程仍然采用传统翻译教学方法,教学效果不佳,培养的人才实际翻译工作能力不强,并不能有效地达到翻译专业的培养目标。^[2]

当今世界,随着大数据时代的到来,数据已经渗透到各行各业,如何有效地利用数据驱动学生进行翻译实践和反思,实现“启发式”“探索式”教学与学习,这是目前翻译专业教师们亟待思索的重要问题。

二、本科翻译专业课程教学面临的问题和前景

20世纪80年代以来,全国高校顺应形势,纷纷开设翻译课程,并逐渐形成了“教师理论讲授+举例分析,学生翻译练习+教师评述”的传统翻译教学模式。这种传统的教学模式培养了一大批优秀翻译人才,但是随着社会分工日益细化,信息技术迅速发展,知识爆炸性增长,传统的教学模式已然不能适应新形势的需要。

穆雷指出:“中国的笔译教学比较混乱,缺少系统合理的教学手段,教师之间也缺乏交流沟通,其结果是翻译教学效率低下,一些原本应该提高翻译能力的课程被上成了熟练语言能力的后续课程,教师疲惫不堪却收效甚微,学生也觉得兴趣不大……”^[3]教学手段不合理、缺乏系统性,是传统教学模式的一大问题,除此之外,传统翻译课堂节奏缓慢,知识容量少,以教师为中心,学生被动学习,这在如“会展翻译”这类翻译专业笔译板块课程上体现得尤为突出。

Johns 指出:“数据驱动学习”(Data-Driven Learning,以下简称DDL)会引起师生之间传统角色的变化,教师成为研究指导者和合作者,不再是知识的传授者了。”^[4]如此一来,学生就成为翻译课堂中的研究者,“在自主研究所提供的语言数据中分析现象,找到规律,而不是一味去模仿前人的翻译”^[5]。对于“会展翻译”课程而言,课堂教学呈现

收稿日期:2016-11-16

基金项目:重庆第二师范学院教改项目“本科职业翻译能力研究与实践”(JG2016224)

作者简介:梁爽(1983—),女,重庆市人,讲师,研究方向:翻译理论与教学。

出复合型的知识结构和技能训练结构,既要涉及翻译技能的提高,又要涉及行业知识,单靠教师讲授和课堂练习远远不能满足教学的需要,很难达到教学目标。因此,学生通过海量的语言数据进行研究,总结经验并找出规律,在“启发”和“探究”中学习,会取得更大的成效。

三、语料库数据驱动

(一) 数据驱动学习理论

数据驱动学习(DDL)是一种基于语料库数据的外语学习方法,自 Tim Johns 于 1991 年提出以来,就引起了国内外教育研究者的广泛关注,给传统的语言教学方法带来了巨大的冲击。根据该理论,“语言教师的任务是为学生者提供语境,让他们能够形成发现式学习的策略,学会自主学习”。^[6]学习者逐渐变成研究者,通过观察大量的语料库数据,归纳总结语言的使用规律,自行发现词语搭配现象、语法规则、语用特征和文本风格,而教师则化身为学生学习的合作者及引导者,激发学生的学习兴趣,促进学生的“探索式”与“发现式”自主学习。

(二) 语料库与翻译教学

虽然数据驱动学习最开始应用于外语教学,但是近年来随着机读语料库的迅猛发展,翻译教学的应用模式得到了极大的改变。Bernardini 建议,翻译教学应辅以平行语料库检索,使翻译专业学生形成一种翻译的“意识”、“反射”和“应变”,这些技能“使专业译员有别于那些不熟练的业余爱好者”。^[7-8]

语料库极大地扩展了语言素材,克服了教师在实际翻译教学中出现的知识储备不足和语料匮乏等问题,让学习者能够更加直观地获得真实自然的材料,带着问题到语料库中寻找答案,而不是通过教师自省式的总结得到单一的参考译文。Bowker 指出,词语索引工具在翻译中的主要优势就是能将包含译者所需词语的文本同时罗列出来,让译者从中发现不同的语言和概念模式,然而这在单一的纸质语料中常常是很难发现的。^[6]

虽然国内对语料库的开发如火如荼地展开,如北京外国语大学中国外语教育研究中心近年建成了 3000 万字词的通用汉英平行语料库,但是对于课堂教学而言,目前还很难找到一个简便、稳定、可下载且经济实惠的汉英/英汉平行语料库。王克非、秦洪武指出:“10 年过去了,双语语料库在翻译教学过程如何有效运用依旧存在许多问题。其核心问题是平行语料库如何与微观翻译教学环节有效衔接。”^[9]对于会展翻译教学而言,更是缺乏一个可靠的会展

类平行语料库。因此,笔者在课堂教学中尝试采用单语语料库辅助翻译实践。相对于平行语料库而言,单语语料库规模更为庞大,数据更为详尽,语料自然真实,使用更加方便,因此是翻译教师应该加以充分利用的有效工具。

(三) 单语语料库

在课堂教学中,网络首先就是最大的一个语料库。Boulton 指出:“目前众多语料库的界面允许用户将整个网络用做语料库。”^[10]除此之外,教学还会频繁使用目前全球两大英语免费单语语料库——COCA 和 BNC。

COCA 即美国当代英语语料库,是迄今规模最大的英语语料库,由美国 Brigham Young University 开发,现有库容 4.5 亿词,涵盖美国 1990—2015 年口语、小说、流行杂志、报纸和学术期刊五大类型的语料,并且每一年中这五方面的语料都是按比例均衡分布的。

BNC 为英国国家语料库,是目前世界上最具代表性的当代英语语料库之一,由英国牛津出版社、朗文出版公司、牛津大学计算机服务中心、兰卡斯特大学英语计算机中心以及大英图书馆等联合开发建立,于 1994 年完成。英国国家语料库词容量超过一亿,由 4124 篇代表广泛的现代英式英语文本构成。

这两大语料库权威性很高,且使用方便,完全免费,是辅助翻译教学的利器。

四、利用网络搜索与单语语料库开展会展翻译 DDL 教学

(一) 利用网络搜索获取翔实语料促进会展翻译教学

笔者曾在国内最大的购书平台当当网上进行搜索,只找到两本会展翻译类书籍——《会展翻译研究与实践》及《会展英语翻译》。前者偏重理论,而后者专注于英汉翻译,语料过于简单。因此课程所需要的大量语料还只能另行收集。如在会展申报文案翻译环节,北京奥运会和上海世博会的申请报告中文版本可以在网络上搜到,还能搜到英文对照版实体书的购买地址。在展台搭建规则部分,可通过搜索关键词“标准展位”“展位类型”“standard booth”“scheme shell booth”“booth type”等找到相关中英文网站,从中选择较新的可靠语料用以课堂范例讲解和翻译练习。

虽然没有专门的会展语料库,但是该课程以整个互联网作为最大的语料库进行语料收集。教师与

学生可利用收集到的可靠语料共建小型会展语料库,对未来的教学能起到辅助作用,也能用于研究目的。

(二) 利用单语语料库提升学习者翻译准确性

单语语料库资源丰富,内容详尽,且搜索功能强大。学习者可以使用翻译材料中的关键词在单语语料库中进行搜索,找到相似的文本,从宏观和微观层面识别用法信息,促使学习者自主研究和反思,能够有效地提升学习者翻译准确性。在会展翻译课程中,教师采用 DDL 教学模式引导学生自主完成会展申报文案、会展指南、会展中心简介等相关材料的汉英翻译任务。课程并无事先指定的教材,教学材料均由教师一手策划。

例如,在某次会展指南翻译任务课上,教师先把学生分为 4~5 人的小组,然后向每组布置相同的汉译英任务。

翻译材料:重庆江北国际机场位于中国重庆市渝北区两路街道,距离市中心 19 公里,是西南地区三大枢纽机场之一。重庆江北国际机场国内定期直飞航班每周 3,487 架次,国际定期直飞航班每周 468 架次。2015 年,重庆江北国际机场完成旅客吞吐量 32,402,196 人次,比上年增长 10.72%;2015 年上半年,机场完成旅客吞吐量 1,562 万人次,旅客吞吐量在中国八大区域性枢纽机场中排第 4 位,比上年增长 11.4%,增速位居中国八大机场之首。

在学生通读材料之后,教师要求学生找出搜索关键词,在网络和 Wikipedia 中搜索内容相似的英文语料;要求学生在找到相关语料后,将之与翻译材料的内容进行对比,从字词层面上升至句子段落层面,认真分析,进行仿写。每个小组必须在规定时间内以小组为单位上交一份译文,并提交翻译报告,解释自己如何利用单语语料库解决翻译难点。由于篇幅有限,笔者将着重讲解学生如何利用语料解决“开通多少国内外航班”与“旅客吞吐量”两个问题。

学生小组均将“airport”定为搜索关键词,随即多方搜集英文语料。一个小组在 Wikipedia 中找到

“Beijing Capital International Airport”以及“Heathrow Airport”两份语料,并将之复制粘贴到 word 文档上,并命名为“airport”,以便进行词语搜索。

S1: 重庆江北国际机场国内定期航班每周 3,487 架次,国际定期航班每周 468 架次。

初译: There are 3,487 regular domestic flights and 468 regular international flights every week in the Chongqing Jiangbei International Airport.

在初译稿中,该小组采用了最常见的“there be”句型。小组成员在讨论后认为,虽然该译文语法正确,但是表述专业性不强,语篇结构相当松散。随后,他们以“domestic flight”和“international flight”作为关键词,在语料中进行查找,搜索到相关用法若干,以下为两个最典型的例句:

例 1: September 2004 saw the opening of a renovated Terminal 1, which at that time solely handled China Southern Airlines' domestic and international flights from Beijing.

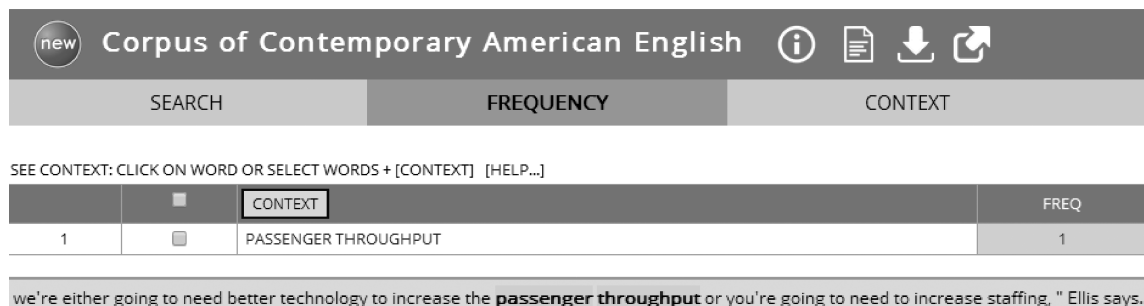
例 2: Prior to the opening of Terminal 3, all international flights (and the majority of the domestic flights) operated from this terminal.

在搜集的相关句子中,“handle”和“operate”使用频率很高。该小组经过分析,认为应突出“airport”作为主语,而“handle”则是最合适的谓语。

因此,本句的译文修改为: The Chongqing Jiangbei International Airport handles 3,487 regular domestic flights and 468 regular international flights every week.

S2: 旅客吞吐量在中国八大区域性枢纽机场中排第 4 位。

显然“旅客吞吐量”是专业术语,学生小组首先求助网络来解决术语翻译问题,百度翻译的译文是“passenger throughput”,为了印证该译文的准确性,学生小组将之作为关键词在 COCA 中进行查找,查找结果如图 1。



The screenshot shows the COCA search interface. At the top, there's a header with 'new', 'Corpus of Contemporary American English', and icons for search, frequency, context, and download. Below the header, there are three tabs: 'SEARCH', 'FREQUENCY', and 'CONTEXT'. The 'CONTEXT' tab is selected. Below the tabs, there's a table with columns for 'SEE CONTEXT: CLICK ON WORD OR SELECT WORDS + [CONTEXT] [HELP...]', 'CONTEXT', and 'FREQ'. The table shows one result for 'passenger throughput' with a frequency of 1. Below the table, there's a snippet of text: 'we're either going to need better technology to increase the passenger throughput or you're going to need to increase staffing.' Ellis says.

SEE CONTEXT: CLICK ON WORD OR SELECT WORDS + [CONTEXT] [HELP...]	CONTEXT	FREQ
1	PASSENGER THROUGHPUT	1

we're either going to need better technology to increase the **passenger throughput** or you're going to need to increase staffing." Ellis says.

图 1 “passenger throughput”在 COCA 中的查找结果

继而又在 BNC 中进行查找,查找结果为“*There are no matching records*”。

由此可见,虽然“*passenger throughput*”也有使用,但是使用频率偏低。随后,学生小组又在“*airport*”文档中查找,发现“*passenger throughput*”并未出现。因此,他们认为应该重新找寻更为合适的英文术语来表达“旅客吞吐量”。他们将“*passenger*”定为关键词,在“*airport*”文档中进行搜索,找到许多包含该关键词的句子,在经过筛选之后,保留了若干相关语料,以下举两例为证。

例3: *Heathrow is the busiest airport in the United Kingdom, as well as the busiest airport in Europe by passenger traffic, and sixth busiest airport in the world by total passenger traffic.*

例4: *It had become the busiest airport in Asia in terms of passenger traffic and total traffic movements by 2009.*

在搜索到的相关语料中,“*passenger traffic*”显然表示了“旅客流量”,而“旅客吞吐量”的意思是“报告期内经由水路、航空等乘船(飞机)进、出港区范围的旅客数量”,因此这个表达从语义而言是准确的。随后,学生们在 COCA 和 BNC 中进行查找,查找结果分别有 99 条和 66 条。这说明,“*passenger traffic*”是广泛的用法。根据语料,该小组将译文定为: *It is the 4th busiest major regional airport in China by total passenger traffic.*

在传统翻译课堂上,学生们都是等待教师给出翻译结果,学习被动,热情不高,效果不佳。然而在 DDL 翻译教学模式之下,学生们尝试自主解决微观翻译问题,或者利用免费的语料库总结语料,分析语义,最终解决翻译问题,既提高了翻译的准确性,也提升了翻译职业能力,由此获得了极大的成就感。

(三) 利用单语语料库让学习者了解语料库的局限性

在某次会展指南(*exhibitor manual*)翻译实践课上,学生们在处理“会展基本信息”时遇到了难题:重庆出租车起步价(含 3 公里)人民币 10 元,超过 3 公里后,每公里运价人民币 2 元。

学生们纷纷表示不知道“起步价”的翻译方式。通过查询百度翻译,学生们找到了两种译法,第一种是“*flag-fall/down price*”,而另一种是“*starting price/fee*”。为了印证两种译法的准确性,学生们自觉地打开了 BNC 和 COCA 的界面,输入关键词进行查询。结果显示,“*starting fee*”在 BNC 和 COCA 中的

使用频率为零,但是“*starting price*”在两个单语语料库中的使用频率都较高。根据查询结果,学生们去掉了“*fee*”,保留下“*price*”。随后,他们在界面上点开“*CONTEXT*”,仔细分析出现“*starting price*”的语料,发现该词应用范围广泛,并不仅仅是用于表述出租车的起步价。然而在 BNC 和 COCA 中,“*flag-fall/down price*”的使用频率却为零。基于这种查询结果,学生们就选定了“*starting price*”,认为“*flag-fall/down price*”为错误表达。

笔者随即向学生们提出问题: *flag-fall price* 这种“错误”用法的来源是什么? 学生们马上转而使用 Wikipedia 和百度进行查找。根据 Wikipedia,“*flag fall*”被解释为“*Flagfall or flag fall is a common Australian expression for a fixed start fee, especially in the taxi and haulage and railroad industries.*”由此可见, *flag-fall price* 用作表述出租车的起步价是澳洲英语表达方式。

也有学生在搜索过程中,收集到了一份来自 CNN 对本国出租车的新闻报告,其中提到:“*Washington (CNN) — Passengers hailing cabs in Chicago starting Sunday may want to read the fine print on the taxi placards before they get in. In addition to the base fare, cab drivers for the first time can charge sick passengers a MYM50 vomit cleanup fee.*”从这份语料中,笔者与学生们发现,美国英语中会使用“*base fare*”来表述出租车的起步价。

在 DDL 教学模式之下,学生们发现单语语料库收集语料的局限性,并通过广泛的语料查找和分析,丰富了语料,解决了翻译问题。学生们认识到,语料库只是提高翻译准确性的辅助工具,还得通过自己的主动学习来积累和丰富语料,并且提升自己对于英美用语差异的意识。

五、结语

数据驱动学习是基于教育技术发展背景之下的新型教学模式,能够给翻译教学带来全新的突破。实践证明,数据驱动学习应用于会展翻译教学具有如下几大优势:第一,改变传统翻译课堂上教师一言堂的状况,教师成为启发者和合作者,而学生成为探索者和研究者;第二,教师学生自行收集会展语料,自建小型语料库,丰富教学材料,从中分析语言现象,解决微观语言问题,由此提升学生的参与感和成就感;第三,通过与大型权威单语语料库进行对比,发现翻译中存在的问题,提高翻译的准确性。

但是,数据驱动学习应用于会展翻译教学也有

一定的局限性:第一,网络上搜索到的语料良莠不齐,需要花费大量时间精力来判断其准确性,否则会造成严重的翻译错误;第二,单语语料库中常常无法检索到会展类专业词汇、句法或者语篇;第三,从译者的角度来看,译者在跨语言、文化的翻译活动中,不仅必须具备双语能力,还必须具备转换能力,而这些能力的培养绝不仅是会使用语料库就能替代的。

由此可见,数据驱动学习在会展翻译教学中只是一种重要的教学模式,并不能给翻译学习带来翻天覆地的变化。我们应该尽可能利用其优势,转变学习者在翻译学习中的态度,引导他们积极主动进行探索,一点一滴地提高自己的翻译能力和职业能力。

参考文献:

- [1] 俞华,朱立文.会展学原理[M].北京:机械工业出版社,2005.
- [2] 罗选民.中国的翻译教学:问题与前景[J].中国翻译,2002(4):56-58.
- [3] 穆雷.翻译教学发展的途径[J].中国翻译,2004(5):25-

26.

- [4] Johns, Tim. 1991. Should You be Persuaded: Two Examples of Data-driven Learning [M] // Johns T., King, P. Classroom Concordancing. Birmingham: University of Birmingham, 1991:1-13.
- [5] 朱晓敏.基于 COCA 语料库和 CCL 语料库的翻译教学探索[J].外语教学理论与实践,2011(1):32-37.
- [6] Lynne Bowker. Towards a Methodology for a Corpus-based Approach to Translation Evaluation [J]. Meta, 2001(2): 345-364.
- [7] 王克非.双语平行语料库在翻译教学上的用途[J].外语电化教学,2004(6):27-32.
- [8] Hunston, S. Corpora in Applied Linguistics [M]. Cambridge: Cambridge University Press, 2002.
- [9] 王克非,秦洪武.论平行语料库在翻译教学中的应用[J].外语教学与研究,2015(5):763-771.
- [10] Alex Boulton. Data-driven Learning: on Paper, in Practice [M] // T. Harris & M. Moreno Jaen (eds). Corpus Linguistics in Language Teaching, 2009:17-52.

[责任编辑 亦 筱]

(上接第 112 页)

误加、误代和错序。偏误产生的原因是:对现代汉语丰富的约量表达方式习得不全面,造成约量表达选择单一;对“上下、前后、左右”限制对象认知不清,致使错序出现;对“上下、前后、左右”三者限制对象的差异辨别不清,造成彼此误代。针对汉语学习者习得约量义对举方位词“上下、前后、左右”的情况,本文提出了三条相应的教学策略:约量表达的个别教学与系统性教学相结合;分解学习难点、降低学习难度,循序渐进;直观性教学与认知引导相结合。

注释:

- ①尽管在北京语言大学 HSK 动态作文语料库中,“X + 前后”只出现了一个错句,但这并不能否认这种错误的系统性,因为在其他语料中有类似的错句出现。

- ②本文所有偏误语料均取自北京语言大学 HSK 动态作文语料库。

参考文献:

- [1] 张豫峰.“x + 前后/左右/上下”的分析[J].语言教学与研究,2004(3):30-36.
- [2] 牛顺心.对举的方位复合词[J].邵阳师范高等专科学校学报,2004(1):108-112.
- [3] 宋红梅.对称式方位复合词“前后/上下/左右”的认知研究[J].长春大学学报,2009(11):39-42.
- [4] 马喆,邵敬敏.论反义方位复合词的约量表达[J].学术研究,2009(5):149-153.
- [5] 钱彬,沈君莉.表估量“X + 上下/前后/左右”的考察[J].现代语文(语言研究),2015(1):68-70.

[责任编辑 亦 筱]