

数据驱动学习方法对大学生英语 写作水平的影响

倪修璟

(上海理工大学 外语学院, 上海 200093)

摘要: 数据驱动学习方法是一种基于语料库的学习方法。在教师的引导之下, 学生运用数据驱动学习方法可以自主学习词汇的搭配知识和语法知识。根据数据驱动学习的理论基础及其特点, 在上海理工大学2014级部分非英语专业大学英语新生中设计并实施利用语料库进行在线自主学习的英语写作教学实验, 实验结果显示, 数据驱动学习方法的确能大幅度提高学生作文的成绩。

关键词: 数据驱动学习; 自主学习; 词汇搭配; 英语写作

中图分类号: H 319 **文献标志码:** A **文章编号:** 1009-895X(2016)03-0267-05

DOI: 10.13256/j.cnki.jusst.sse.2016.03.013

Effects of DDL Method on the English Writing of Non-English Majors

Ni Xiujing

(College of Foreign Languages, University of Shanghai for Science and Technology, Shanghai 200093, China)

Abstract: Data Driven Learning (DDL) is a corpus-based learning method, which can help students, under the guidance of the teacher, to carry out autonomous learning on the English collocation and grammar. The author illustrates the theoretical basis and characteristics of DDL, designs and implements a classroom experiment in which the subjects from part of non-English major freshmen (Grade 2014) studying College English Band 1 adopt the method of autonomous learning with the help of online corpora. The results of the experiment show that Data Driven Learning may have significantly raised the students' scores in English composition.

Keywords: DDL; autonomous learning; vocabulary collocation; English writing

一、研究背景

数据驱动学习 (Data Driven Learning) 是一种基于语料库的学习方法。该方法是由 Tim Johns 教授在 20 世纪 90 年代初提出的。目前, 国内外数据驱动学习的研究已经取得了一定成果^[1-2], 但侧重

于理论和技术的介绍, 相关的课堂应用研究并不多见^[3]。张北镇等对间接数据驱动学习的课堂实现模式进行了实证研究^[4]。张玲的《基于数据驱动学习的机辅写作教学模式初探》研究缺乏实证数据^[5]。鉴于此, 本文拟对数据驱动学习方法在英语写作的应用进行实证研究。

收稿日期: 2015-09-06

作者简介: 倪修璟 (1964-), 男, 副教授。研究方向: 语料库语言学、英汉翻译。E-mail: nixiujing@126.com

二、数据驱动学习方法的理论基础及其特点

Johns 教授认为,利用数据驱动学习,学习者既是研究者,又是侦探夏洛克·福尔摩斯。数据驱动学习的基本理念符合“以学习者为中心”的思想,即教学活动以学习者自主学习为主^[6]。可以说,数据驱动学习方法是建构主义为理论基础的,因为建构主义理论提倡在教师指导下以学习者为中心的学习,强调学习者对知识的主动探索及对所学知识意义的主动构建。数据驱动学习在很多情况下都离不开网络和计算机。在外语学习中,利用在线语料库,可以发现某一词汇的典型搭配词、使用频率、语体及语义韵等信息,甚至可以完成某一科研任务。当学习者能够主动利用外部的学习环境和学习资源,自我管控学习方式、内容和过程时,学习者便实施了某种自主学习的活动^[7]。在传统的英语课堂教学模式中,教师是整个教与学过程的中心、教学的主导者、知识的传授者,学生是外部刺激的被动接受者^[8]。数据驱动学习模式下,Johns 教授鼓励学生积极主动地从真实语料中去发现语言事实。事实上,数据驱动学习不是直接教授某一语言特征,而是向学生提供语言事实,引导学生基于大量的语料库数据观察、概括和归纳语言使用现象,自我发现语法规则、意义表达及语用特征,学生据此提出假设并最终得出结论。有两种类型的数据驱动学习:直接数据驱动学习(Hands-on DDL)和课堂数据驱动学习

(Hands-off DDL)。前者就是学习者直接与计算机接触,亲自动手对某一语料库进行检索。后者就是学习者不必直接接触计算机,语料库的检索工作由教师代之完成;教师根据教学需要,对检索结果进行选择,分类,编制相应的学习材料,供学生课堂使用。本研究涉及的是直接数据驱动学习,文中简称为数据驱动学习。同传统外语教学相比,数据驱动外语学习具有以下主要特征:减少对教师的依赖,以学生的自主学习为主要的过程;以真实语言为主要语言输入^[9];强调探索和发现的学习过程;主张自下而上,归纳式的学习^[10-12]。学生从语料库中挖掘到某些数据之后,往往会发现被教师所忽视的一些语言现象,学生们也可能发现教科书中没有提到的语言现象,从而发挥他们的学习主动性,使他们通过语料库学习到更多的语言知识,思考更多的语言问题。数据驱动学习又称为发现式学习,是由教师向学生提出一个需要解决的问题,在教师的指导下,学生利用海量的语料库数据,运用检索软件,立地搜索、分析、处理语料,从中发现和归纳语言规律,从而达到解决某一问题的目的^[13]。使用数据驱动学习方法,学生可以在极短的时间内,反复接触到语境中的关键词(KWIC),快速查找出某个节点词的搭配词或推断出搭配词的某些共同的语义特征。通常为了让学生尽快掌握数据驱动学习方法,建议教师开始教学时使用课堂数据驱动学习方法。例如,在英国国家语料库中,人们可以检索到 owing to 的 KWIC 索引行句子。学生在教师的引导下,可以发现 owing to 搭配词的语义特征。

表1 owing to 的 KWIC 索引行例句
Tab.1 KWIC examples of owing to

左跨距	节点词	右跨距
the war and were again humiliated at Wembley in 1966.	owing to	a bizarre mix-up between a Japanese morse operator
the comedy reaches its height as it emerges that,	owing to	a design oversight,ED 209 can not negotiate stairs.
and Antiquities of the County of Dorset are very rare	owing to	a disastrous fire. According to the Bronte sisters,
the work of building the circle was never completed and	owing to	a miscalculation or perhaps an unfavourable omen,
Occasionally tours may be cancelled at short notice	owing to	circumstances beyond our control. If possible an
Capt. Warr and party,all very tired and fatigued	owing to	lack of food and water for four days. We set

也可以发现与 totally 搭配的一些高频形容词,如 different, new, dependent, wrong, unacceptable, inadequate, unexpected, unsuitable, unaware, independent, 等等,大多具有消极或中性语义韵。为了更明确地向

学生说明数据驱动学习方法在英语词汇学习中的作用,课堂教学可选用更大容量、更新的美国当代英语语料库的语料,采用比表1更为完整的 KWIC 索引行例句。

学生对英语词汇搭配的概念往往似懂非懂,但是,他们是否可以在教师的引导之下,借助于具有搜索引擎功能的在线语料库,去发现正确的搭配词,并且应用于作文的写作和批改呢?为此,本文进行了一个利用数据驱动学习进行英语写作的实证研究。

三、实验设计

(一) 实验目的

本实验目的是,实验班学生在教师的引导下,采用数据驱动学习法,利用 BNC、COCA 等在线语料库检索工具检索指定词和作文中的搭配不当的词,通过检索节点词和搭配词的语体特征、使用频率及语义韵,探究在线自主学习对学生自改作文的影响。实验假设是:实验对象通过在线语料库自主学习写作,包括在线提交及教师引导下的学生自主修改等活动,能明显提高实验对象的作文成绩。

(二) 实验对象

参加本实验的学生为上海理工大学 2014 级的 21 名新生,另有 21 名为对照组。这 42 名学生高考成绩接近,之前均未接触过语料库,也没有数据驱动学习的经历。实验前对实验组和对照组的学生进行作文前测。该测验要求实验班和对照班的学生完成 150 字左右的作文。

(三) 学习材料及要求

教师从英美报刊杂志选取包含下列词汇的 16 篇短文(每篇包含一个): thanks to, owing to, obtain, gain, effect, influence, identify, recognize, challenging, difficult, neglect, ignore, contribute, donate, environment, surroundings。这 16 个词汇组成 8 对近义词(组),对学生来说,通常有一定的困难。实验组和对照组的学生均需要学习这 16 篇短文及这 8 组近义词的区别,课堂教学需完成词汇辨析及搭配等练习。并且在课后完成一个短学期 5 篇短文的写作练习。同一教师对实验组和对照组的学生讲解这 16 篇短文,不同的是,对照组学生所学有关近义词的内容基本上来自于英语同义词词典,或朗文、或牛津等英语工具书。而实验组的学生需要首先学会语料库、检索方法及数据驱动学习方法,亲自动手检索 BNC 和 COCA 语料库,自己归纳总结近义词的异同。

(四) 实验步骤

首先,为了让受试学生尽快适应数据驱动学习方法,教师将事先挑选编辑好的美国当代英语语料库 COCA “关键词语境”(KWIC)完整例句打印出来发给学生,或通过多媒体大屏幕展现给学生,引导学生观察,思考,提出假设,验证假设,进而总结某词搭配词规律性的东西。例如,受试学生可通过观察短语 thanks to 的 KWIC 例句,归纳出 thanks to 搭配词的语义特征或语义韵。

表 2 thanks to 的 KWIC 索引行例句

Tab.2 KWIC examples of thanks to

左跨距	节点词	右跨距
a ton of hope about where this little guy can go in life.	thanks to	the remarkable work of his doctors and nurses
however, today's watcher can be virtually all-seeing,	thanks to	GPS monitoring technology.
about our preferences long before we're dying.	thanks to	medical advances, people today live longer than at any time
this pitfall..., has so far been largely avoided	thanks to	the presence of a number of high-level specialists...
travel is an extremely safe mode of transportation and,	thanks to	improvements in technology, is getting safer.
they should be up more,	thanks to	the addition of more dealers and another model

受试学生在多媒体教室接受为期八周的训练,每人配备一台联网的计算机,每次上课时间 90 分钟,教学内容为一组近义词。在教师的指导下,采用数据驱动学习法检索在线语料库 BNC 和 COCA,第二周至第六周学习指定词汇的高频搭配词及其语义特征。同时,每周课外完成一篇短文练习。要求受试学生学会用数据驱动学习方法辨别所学词汇,并解决写作过程中遇到的疑难问题。受试学生以电

子文本的格式发邮件上交给教师。教师批改时,不对学生的作文进行直接修改,仅仅指出作文中的用词错误、语法错误及搭配不尽如人意的地方,发现每个实验学生的作文都存在这样或那样的问题。一般情况下,学生能通过查找在线语料库,有时借助于句酷批改网,解决大部分搭配问题和一部分语法问题。有些学生的作文自己改两至三次,才能基本达到教师的要求。第九周实验组和对照组均参加短

学期作文测试,后测作文题材与前测相同。

对于对照班作文的提交及批改,教师采用传统的方法:学生上交纸质版的作文,教师指出作文的优缺点,改正作文中的病句及用词错误。但是,两个班评分都是按照大学英语四级作文的评分标准进行,学生的实际得分由15分制转换为百分制。

四、实验结果与分析

实验前后的作文成绩均由批改网判定,并且采用SPSS统计软件进行分析,结果如表3和表4。

表3 实验班和对照班作文前后测 means 描述统计

Tab.3 Results of descriptive statistics of pretest and posttest essays scores of experimental and control groups

项目	班级	人数	均值	标准差	均值标准误差
前测	1	21	66.57	5.016	1.095
作文分数	2	21	66.81	5.671	1.238
后测	1	21	77.26	4.008	0.875
作文分数	2	21	71.81	3.916	0.855

表4 实验班和对照班实验前后独立样本 T 检验

Tab.4 Independent samples T-test analysis results of pretest and posttest essays scores of experimental and control groups

阶段		方差齐性检验			均值 t 检验				
		f 值	t 值	自由度	显著性	平均差	平均差标准误差	95% 置信区间	
								Lower	Upper
前测	EVA	0.051	0.144	40.000	0.886	0.238	1.652	-3.577	3.101
	EVNA	0.822	0.144	39.410	0.886	0.238	1.652	-3.579	3.103
后测	EVA	0.108	-4.459	40.000	0.000	-5.452	1.223	-2.981	-7.924
	EVNA	0.744	-4.459	39.978	0.000	-5.452	1.223	-2.981	-7.924

注:EVA = equal variance assumed;EVNA = equal variance assumed

实验前测中,实验班(Class 1)平均分66.57与对照班66.81无显著差异,独立样本T检验结果($t = -0.144$, $df = 40$, $p = 0.822 > 0.05$),未达到统计学上的显著意义;而实验后测显示实验班平均分77.26明显高于对照班平均分71.81(表3),独立样本T检验结果 $t = -4.459$, $df = 40$, $p = 0.000 < 0.05$,两个班平均相差5.45分(表4),有显著差异,即,实验班的作文后测成绩明显高于对照班。因此,不同的教学方法对大学生英语写作水平产生了不同影响。

实验班学生在教师的引导下,通过语料库在线自主学习英语作文中某些词汇的搭配,英文词汇习得效果有了显著的提高。教师指出不当搭配后,实验班学生可以立即应用数据驱动学习所获得的词汇搭配知识用来修改自己的作文。当然,利用计算机或语料库,学生也可轻易发现语法错误,在语料库中迅速推断出某词的语法规则。而对照班的学生平时不会采用数据驱动学习方法,往往可能会孤立地记忆单词或语法规则,注意不到语块的学习,从而会影响语言表达的准确程度和流利程度。语料库的语料是一种浓缩的阅读,学生能从中概括有用的句型及用法,并且有效地把被动词汇转变为积极词汇^[14]。

五、数据驱动学习的优势及其与 ESP 之关系

就写作而言,数据驱动学习方法首先能在极短时间内便捷地给学习者提供有关词汇的语体特征、语义韵特征、使用频率及其搭配词的相关信息,而利用传统的教学方法却很难获悉这方面的信息。词典和教科书不易在简短的篇幅内阐释复杂的语体特征,而数据驱动学习方法会帮助学习者轻而易举地解决语体等问题。这样就极大地方便学生作文中正确选词。数据驱动学习方法的第二个优势在于培养学生的发现式学习能力,这是传统的填鸭式教学无法做到的,这种学习方法能培养学生的创新思维能力。第三,ESP的教学或科研在某些方面或将离不开数据驱动学习。在国内学术界,卢思源先生真正可以算得上是ESP的奠基人,早在20世纪70年代末,卢先生所任教的上海机械学院在全国首创科技英语专业。这在当时是轰动全国的课程设置大事件,其影响之深远至少可以从后来中央电视台连播卢思源先生主持的工程师英语讲座看出。然而,近四十

年后,传统 ESP 词汇教学及非英语专业学生用英语所写的专业论文中词汇的使用还不尽如人意,有些问题是可以数据驱动教学模式解决的,但是,这方面的研究还比较薄弱。如今,在国内相当多的学者提出大学英语应该向 ESP 转型的背景下,提倡应用数据驱动学习方法的英语论文写作或词汇教学不仅必要,而且具有可行性。可以预见,在不久的将来,该方法很有可能发展为 ESP 词汇教学和写作教学的一种主流方法。

六、结束语

据后测后对学生的访谈可知,学生都学会了自己观察、分析真实语境资料并且培养归纳词汇用法的能力,这种发现式的学习方法的确提高了学生学习的积极性和自觉性。虽然对照班后测作文成绩也有提高,但是远远落后于实验班。采用数据驱动学习方法指导学生英语作文能大幅提高学生作文成绩,可能是因为在此过程中,学生可以接触大量的真实语料,通过自己的观察,便捷地分析归纳词汇的用法,证实自己的假设,从而产生学习中的成就感。数据驱动学习方法的确可以鼓励学生使用语料库及相关检索工具,寻找正确的词汇搭配^[15]。数据驱动学习方法不仅充分挖掘了学生的学习潜能,而且调动了他们的学习积极性。虽然自主学习和学习成绩之间可能不是一种简单的因果关系,但是近年来,以教师和教材为中心的传统外语教学理念已经受到了挑战^[16]。随着计算机、语料库语言学及网络技术的发展,传统的教师传授学生接受的“填鸭式”教学模式正在受到来自“多媒体、多模态、多情境、多资源渠道的综合立体、开放式学习系统”的种种挑战。信息技术与课程整合的研究,已经使外语教学形成了新的教育信息化教学范式^[17]。传统的教学模式“被教师引导的、学生积极参与的课堂教学和课外自主式学习的复合型学习模式所取代已成为必然”^[18]。

参考文献:

- [1] Hunston S. Corpora in Applied Linguistics [M]. Cambridge: Cambridge University Press, 2002.
- [2] 俞燕明. 数据驱动词汇教学——基于计算机和语料库的研究性教学探索[J]. 外语电化教学, 2009(3): 58-62.
- [3] 朱慧敏. 数据驱动学习: 英语词汇教学的新趋势[J]. 外语电化教学, 2011(1): 46-50, 59.
- [4] 张北镇, 周江林. 数据驱动学习的课堂实现模式研究[J]. 外语与外语教学, 2012(3): 41-45.
- [5] 张玲. 基于数据驱动学习的机辅写作教学模式初探[J]. 现代教育技术, 2007, 17(8): 28-33.
- [6] 许家金. 词汇中心教学法的交际观——理论溯源与反思[J]. 中国外语教育, 2009, 2(4): 38-45.
- [7] 王懿, 宣安. 基于网络的交互式自主学习模式对大学生英语写作水平和动机的影响[J]. 外语电化教学, 2010(5): 48-52.
- [8] 陈坚林. 关于“中心”的辨析——兼谈“基于计算机和课堂英语多媒体教学模式”中的“学生中心论”[J]. 外语电化教学, 2005(5): 3-8.
- [9] Clifton J, Phillips D. Ensuring high surrender value for corporate clients and increasing authority of the language instructor: the dividends of a data-driven lexical approach to ESP [J]. Journal of Language for International Business, 2006, 17(2): 72-81.
- [10] Johns T. From printout to handout: grammar and vocabulary teaching in the context of data-driven Learning [J]. English Language Research Journal, 1991, 4: 27-45.
- [11] Kennedy C, Miceli T. An evaluation of intermediate students' approaches to corpus investigation [J]. Language Learning & Technology, 2001, 5(3): 77-90.
- [12] Lee D, Swales J. A corpus-based EAP course for NNS doctoral students: moving from available specialized corpora to self-compiled corpora [J]. English for Specific Purposes, 2006, 25(1): 56-75.
- [13] 倪修璟. 数据驱动学习法在英语词汇教学中的应用研究 [C]//上海市外文学会高职高专外语教学专业委员会. 中国外语教学探索与研究: 理论与实践. 上海: 复旦大学出版社, 2012: 285-291.
- [14] 屈典宁, 邓军. 基于语料库的语块习得模式研究[J]. 外语界, 2010(1): 47-53.
- [15] Chen H J H. Developing and evaluating a web-based collocation retrieval tool for EFL students and teachers [J]. Computer Assisted Language Learning, 2011, 24(1): 59-76.
- [16] 罗凌, 温善毅, 朱永红. 英语语法引导式数据驱动学习平台的构建与成效[J]. 外语电化教学, 2014(5): 16-21.
- [17] 李艳, 胡加圣. 外语教学的范式转换与外语教育技术的学科理据分析[J]. 现代教育技术, 2011, 21(12): 62-66.
- [18] 张殿玉. 英语学习策略与自主学习[J]. 外语教学, 2005, 26(1): 49-55.

(编辑: 朱渭波)