

语料库及数据驱动学习辅助高中英语词汇教学的优势

权晓辉, 吴江

(西北师范大学 外国语学院, 甘肃 兰州 730070)

摘要: 词汇教学作为高中英语教学的一个重要组成部分, 与计算机辅助下的数据驱动学习一起依托于正在蓬勃发展的英语语料库, 从词频统计研究、词语搭配研究、词汇句法层面知识信息的提供以及词汇运用语境、纠正用词失误等5个方面发挥着语料库在高中英语词汇教学领域中的重要作用, 旨在改进词汇教学方法, 提高英语词汇教学的质量。

关键词: 语料库; 英语词汇教学; 优势

中图分类号: H319

1 引言

词汇是语言的建筑材料, 是理解语言和表达语言的基础。词汇也是学习一门语言必备的起点。对于词汇在语言习得中的重要性, Wilkins 有精辟的论断, 他说^[1]: “如果没有语法, 很多东西难以表达; 而没有词汇则什么也表达不了”。同样, Vermeer 也指出词汇知识学习的重要性, 他说道^[2]: “掌握词汇是理解与被理解的关键所在。学习一门新语言的主要任务是掌握词汇, 语法知识不能提升语言能力和流利程度”。在中国的英语学习和教学中, 词汇量少固然是一个问题。但发现许多学生的词汇量尽管很大, 却依旧有交流的困难。如何去学习词汇, 就变成了英语学习和教学的一个重要问题^[3]。Sinclair 在20世纪80年代末就提出外语教学可以从词汇学习入手, 他认为, 在过去词汇教学从属于语法教学, 教师教授的词并不具有代表性和科学性^[4]。在英语教学及学习过程中, 不管是作为英语教学的教师还是英语学习者都认识到英语词汇学习在整个英语学习过程中的重要性。但是, 截至目前, 花费时间多、学习词汇效果差仍然是词汇教与学过程中最突出的问题。造成这种现象的原因有很多, 其中在学校英语教学中没有有效的词汇教授的方法和手段是重要因素。因此本文试图结合目前蓬勃发展的语料库及数据驱动学习探索词汇教学的新途径及它的优势, 以此拓宽视野和思路, 提高高中英语词汇教学效率, 更好地服务于英语词汇教学。

2 词汇的定义

在《牛津高阶英汉词典》中, 词汇被定义为 “一

种语言里所有的(或特定范围的)词和固定短语的总和”。Coady 在1993年总结出什么是“掌握一个单词”的具体含义, 他的总结暗示出在一种语言里掌握词汇的复杂性。他认为对一个单词的掌握应包括以下几方面: 1) 对在不同情况下的一个词和另一个词共现的可能性, 且该词可能同现词的类型; 2) 语域上的限制; 3) 恰当的句法功能; 4) 潜在的词汇形式和派生; 5) 所具有的词汇搭配网络; 6) 词义特征, 引申意义和比喻意义等等^[5]。

鉴于此, 词汇的学习并没有学习其发音和字面意义那么简单。词汇教学和学习应包括下列词汇知识: 搭配、语法特征、隐喻意义、联想意义、情感意义等。对于英语学习者来说, 尤其是急需扩充词汇量的高中英语学生来说, 仅从教科书上教师所教去掌握英语词汇肯定是远远不够的, 脱离英语语言使用在真实语言环境中的参考, 学习者无法真正理解并掌握一个词的具体使用情况。

3 语料库及数据驱动学习

语料库是按照一定的语言原则, 运用随机抽样方法, 收集自然出现的连续的语言, 运用文本或话语片段而建成的具有一定容量的大型电子文本库^[6]。它在一定原则下收集批量的不同种类的口头或笔头语篇素材, 以电子版本的形式储存在电脑中, 用于语言的量化调查和质性分析。它是一种事实性的语言资源, 用来改进人们对语言结构和语言应用的描述并且应用于多个方面, 包括如何用计算机来加工自然语言, 以及如何理解、如何教授语言^[7]。

Tim Johns 在1991年提出数据驱动学习方法 (data-driven learning, 简称 DDL)。其主要思想是

指导外语学习者基于大量语料库语料数据,观察、概括、归纳语言使用现象,从而自我发现并总结出语法规则、意义表达和语用特征^[8]。作为一种语言学习辅助工具,数据驱动学习方法,基于语料库,尤其是基于对语言材料的检索来学习语言的方法,改变传统的教师向学生单向灌输的教学模式,为学习者创建了一个真实的语言环境,能有效地激发学习者的外语学习兴趣。培养他们自主学习和自我解决问题的能力,并达到辅助英语教学的最终目的。

4 语料库及其辅助英语词汇教学的基础

4.1 理论基础

Holec 指出自主学习是学习者对语言学习的自我管理并且负责的能力,即,学习者自己确定学习目标、内容、进度;自己选择学习方法和技巧;自己监控学习过程并且最后自己评价学习效果^[9]。这是对学生学习能力发展的需要,是时代的需要,也是产生新知识的需要,而不是重复旧知识能力^[10]。要授之以“鱼”,更要授之以“渔”。目前,根据我国英语教学和学生自主学习能力的现状来看,让学生完全自主学习并不能够实现,而比较恰当的做法是在课堂教学中,在教师的指导下,有意识地培养学生自主学习能力^[11]。语料库以它丰富真实的语料和独特的检索功能,在数据驱动学习方法下为教师指导下学生进行自主学习提供了一个非常有利的条件。在教学中的应用日益广泛,它以其宏大数据库为基础,涉及词汇大纲和教材编写、词汇教学、语法教学、语篇章分析、错误分析、语言测试及学生自主学习能力培养等诸多方面。

4.2 技术基础——词频和检索

语料库的一个主要贡献就是通过检索增加人们对词频概念的认识。它通过计算机运行相应程序,可以计算、统计出在英语口语或书面语中出现频率最高的英语词汇。因为常用词汇教学的成功与否直接影响着学生掌握英语的程度,这样,我们便能避免片面性,也将大大减少教学的盲目性。

在英语学习过程中,大多数英语学习者习惯将教科书中的词汇及其义项独立,割裂开学习。这样,对于一个词的认识往往是片面且支离破碎的。而在实际语言运用上,词汇与词汇之间在音、形、义方面存在着各种各样的联系。学习者如果孤立地学习一个词汇,那么这种联系就会无法认识了解到。语料库的另一个关键技术——索引有助于解决这个问题

利用现代计算机技术,这种称之为 KWIC(Key - word in Context) 的索引程序来分析处理语料,可以迅速方便地从包含数百万上千万词的语料库中,把某个词或短语在这个语料库中出现的全部实例检索出来,不仅统计出该词或短语出现的频率,而且能更准确更全面地反映词汇之间的关系,体现各语言形式在实际交际中的意义和用法。让语言学习者看到真实的语言使用是如何遵循规则而又不屈从于规则。显然,利用语料库检索技术可以大大提高学习者对词汇的全面理解和应用,而不是句句照足语法。

5 语料库及数据驱动学习在高中英语词汇教学中的应用

在我国,语料库的研究和教学应用开始于20世纪80年代。已建成和在建的有多个语料库,并且种类繁多。有原始语料库、赋码语料库、平行语料库、学习者语料库等。其中在学习者语料库中,比较有代表性的并对我国英语教学有相当作用的是中国学习者英语语料库(CLEC)。CLEC 全称 Chinese Learner English Corpus,即中国学习者英语语料库^[12]。CLEC 的特点是针对中国英语学习者并且有较强的教学支持功能,并对所有的语料进行语法标注和言语失误进行了标注(共61种),借助它分析了中国学生的英语作业有什么特点和失误,显然这对研究中国学习者英语的特点具有很大价值,其成果对推动我国英语教学也具有较高的参考价值。

从目前对语料库的应用和研究来看,词汇研究是最活跃的一个方面。从以下几个方面可以看出语料库及数据驱动学习对高中英语词汇教学的推动作用及其优势。

5.1 词频统计研究

在语料库中,收集和存储了大量不同类型口头或笔头的语言素材,归类、整理和统计这些素材就可以知道某个英语词汇在该语料库中出现的次数,即,频数,这样,就可以计算出该词汇的使用频率。对词频统计研究最直接的应用就是编制词频表,从而确定不同等级高频词汇的范围和数量。因此,基于语料库的词频统计研究可以得出高频词汇。这些高频词是现实生活中的常用词,是人们自然运用语言中最基本、最具活力、最有生产力的词汇。鉴于此,词频统计在英语词汇教学中直接作用在“教什么词”的决策。对于高中英语词汇教学来说,高频词具有非常高的价值,首先,它们为教师和学生提供词汇教学和学习的参考,找出英语词汇教学重点,按主次进

行教学;其次,通过对高频词的学习,学生在高中英语词汇学习中经常遇到他们学习过的词汇,从而获得成就感,建立对英语词汇学习的信心。

5.2 词语搭配研究

利用语料库教授并学习英语词汇搭配,可以在真实的语境下提供丰富的例子,帮助学生理解英语词汇的各种搭配以及该词汇的语法、语用特征。词汇搭配是语料库词汇研究中最活跃的领域,在语料库语言学中处于中心地位。关于它的研究可获得词汇如何使用的一些规则。例如:哪些前置词与特定动词同时出现,哪些动词与有关名词同时出现等。借助 COBUILD 语料库检索 commit 一词,可以随机得到如下检索行:

- 1) Merely by staying on ,did not commit a criminal offence.
- 2) She drove him on to commit murder.
- 3) A knock - out or somebody ' s going to commit a major faux pas and out side of prison walls.
- 4) However ,I feel that by refusing to commit itself to a date for the special.
- 5) The size of the new German army and commit the Allies to synchronized.
- 6) A determined and deliberate act to commit murder.
- 7) In and decide whether he should commit suicide or not.
- 8) Congregation of young people to commit their lives to Christ.
- 9) It doesn ' t matter where he sends me. I didn ' t commit any crime.
- 10) Of a computer keyboard button they commit their companies to high - risk.

由以上检索行,对于中国学生来说,要学习这个不太常用的词 commit 就容易多了。他们可以根据以上检索行总结出 commit 的搭配结构 V + N/PRON + PREP(line 4 ,5 ,8 ,10) ,V + N 两个结构。而且从中可以看出 ,V + N 的出现次数最多。可以发现和 commit 共同出现较多的名词有: offence ,sin ,crime ,harm ,murder ,suicide 等。借助英汉字典,学生可以自己发现以上这些名词全是一些具有消极语义的词汇。通过观察总结,学习者就知道了该词汇如何搭配使用。同样,研究人员和教师也可以从中国学习者语料库中提取该词的使用情况,进行比较研究,从而发现中国英语学习者和英语本族语言使用者在该

词使用方面的差别。以 commit 一词为例,从中国学习者语料库(CLEC) 中提取了该库中所有关于 commit 一词的使用。结果表明,中国学习者对 commit 一词的使用和掌握在以下几方面与英语本族语使用者存在差距。首先,从整个使用来说,中国学习者对该词使用很少,使用率很低。第二,从该词的搭配来看,中国学习者所用到的与 commit 一词共同出现的词只有 crime ,suicide ,而其他的一些词汇几乎很少用到。第三,从结构上来看,中国学习者所使用的结构只有 commit + N 一种。

5.3 提供词汇的句法知识信息

词汇使用也有句法限制,检索语法标注过的语料库,可以了解一个词在句子中能做的成分,并且意思相近的词汇在句法使用中的不同也能体现出来。Douglas 借助语料库分析了 begin 和 start 两个动词的句法限制。在一些学习资料中,这两个词几乎是可以互换使用的。然而,通过语料库的分析才发现并非如此。根据语法常识所说,动词一般要分为及物动词(Transitive) 和不及物动词(Intransitive) , Douglas 通过统计得出该词在使用方面不同的结论: 尽管这两个词在意义和做语法成分上很相似,但这两个词与之在句中实际使用的结构有很大的差异,即它们各自有自己倾向的搭配: start 更多地用于不及物动词,而 begin 一词作及物动词使用,并和 to - clause 搭配使用的情况出现最多,而 start 则除此以外其他结构搭配较多^[13]。

由此可见,那些看似用法意义都相同的同义词、近义词,借助语料库和相关的计算机分析软件,都可以发现它们之间细微的差别。这些差别对于教学和词汇学习是至关重要的。作为英语学习者,了解词汇的具体使用情况,而不仅仅只是在学习该词在母语中对应的意思,更不能把一些在母语中有相同意义的词视为等同。

5.4 提供词汇的运用语境

利用语料库所提供的大量的自然语境中发生的语料,教师可以参考并选取语料库中相关的真实语句作为例句,可以使举例之于呈现环节更加有效,也可以得到目标词汇的各方面信息,如: 高频搭配词项、高频运用义项等。此外,在语料库中搜集的例句用于英语词汇教学时,其语料真实、丰富、新颖,这样,例句本身就会对学习产生很大的吸引力,学习者因此对例句会有较深刻的印象,从而刺激记忆产生良好牢固的记忆效果。总之,利用从语料库的真实语境中提取的例句可以使学习者保持浓厚的学习兴趣,更持久地投入到英语词汇学习中。

5.5 纠正用词失误

学生在词汇应用中经常出现错误,尤其是在写作过程中。语料库(例如上文提到的中国学习者英语语料库)能够提供关于一般学习者的错误状况,教师在备课阶段就可以预先考虑,在上课阶段发现错误时,可以不直接纠正学生出现的错误,而是利用语料库索引软件对使用错误的词项进行索引,使学生通过对正确使用的词汇的索引行的观察分析,达到对用词错误的自我纠正。以这种英语词汇学习方法学习,学习者会对出错词汇的记忆更加深刻,再次使用时也会避免失误,该词使用也会更加恰当。

6 结束语

语料库对语言教学和研究的重要作用已是不可争辩的事实,具体到英语词汇教学的研究与实践领域,它也都正在发挥着巨大的作用。在高中词汇教学中引入语料库,对外语教学思想和方法,都具有非常深刻的意义。这不但是教学形式、教学手段的改变,更是引发了一场教学理念的革新。它改变了我国长期以来以教师为中心的传统教学模式,开始确立了一种以学生为主的积极主动的个性化教学方式。随着对语料库更为广泛深入的研究与应用,我们会更加全面深刻地认识它,在英语词汇教学以及其他众多的研究和实践领域中获得更大的回报,并在更大范围内利用语料库服务于科学的语言研究和教学实践探索。

参考文献:

- [1] Wilkins, D. A. Vocabulary: linguistics in language teaching [M]. London: Edward Arnold Ltd, 1972.
- [2] Vermeer, A. Exploring the second language learner lexicon [A]. In L. Verhoeven and J. H. A. L. De Jong (eds) The

Construct of Language Proficiency [C]. London: Benjamins, 1992, 147-162.

- [3] 褚振莉. 词汇单位的语言学理论模式[J]. 外语教学, 2006, 20(5): 20-22.
- [4] Sinclair J. & Renouf, A. A Lexical Syllabus for Language Learning. In Carter, R. & M. McCarthy. (eds). Vocabulary and Language Teaching. London: Longman, 1988.
- [5] Coady J. Research on ESL/EFL Vocabulary Acquisition: Putting it in Context. In Second Language Reading and Vocabulary Learning [M]. T. Huckin, M. Haaynes, and J. Coady (eds), Norwood, New Jersey: Ablex Publishing Corp, 1993, 3-23.
- [6] 杨慧中. 语料库语言学导论[M]. 上海外语教育出版社, 2002.
- [7] 何安平. 语料库辅助英语教学入门[M]. 北京: 外语教学与研究出版社, 2010.
- [8] Johns T. Should you be persuaded: two examples of data-driven learning? In Johns, T, King, P. (Eds.), Classroom concordancing [A]. Birmingham: ELR, 1991, (a).
- [9] Holec, H. Autonomy and foreign language learning [M]. Oxford: Pergamon Press, 1981.
- [10] Benson, P. & Lor W. Making sense of autonomous language learning: conceptions of learning and readiness for autonomy [J]. English Centre Monograph, No. 2. The University of Hong Kong, 1998, (3).
- [11] 崔岭. 多媒体投影式大学英语课堂教学模式: 学生角色探析[J]. 外语界, 2004, 16(4): 15-17.
- [12] 桂诗春, 杨慧中. 中国学习者英语语料库[C]. 上海外语教育出版社, 2003.
- [13] Douglas B. Conrad, S. & Reppen, R. Corpus Linguistics [M]. Foreign Language Teaching and Research Press, 2002, 8.

(上接第47页)

当然我们也发现,在超低频阶段(即实验结果的第1、2、3阶)两者出现了比较明显的误差,这是因为受扬声器设备的限制,信号发生器输出的超低频信号无法精确的通过扬声器激发出来,致使信号接收器接收的低频共振频率与转移矩阵计算的结果出现误差。这种因实验设备的不足而导致的误差是可以消除的,在今后的实验中可以通过提高实验设备的精度(特别是扬声器的激发精度)来解决。

通过类比法原理,建立了声学模拟实验系统,对实验管道进行了测试,其结果与采用传统的转移矩阵法计算的结果基本吻合。这充分证明该实验测试技术在计算管道气柱固有频率上是可行的、可靠的。

这对研究复杂压缩机管道的振动以及压缩机管道结构的优化设计都具有重要的意义,并且还验证了声学实验方法对压缩机复杂管系气柱固有频率的研究具有借鉴和实验分析价值。

参考文献:

- [1] 姜文全, 杨帆, 王茂延, 等. 基于 ANSYS 的往复式压缩机管系气柱固有频率计算[J]. 压缩机技术, 2008, 36(6): 265-270.
- [2] 刘智勇, 高孟理, 纪燕飞. 活塞式压缩机管路设计中气柱固有频率的计算方法[J]. 化工机械, 2008, 35(4): 212-215.
- [3] 党锡淇, 陈守五. 活塞式压缩机气流脉动与管道振动[M]. 西安交通大学出版社, 1984.