

程 工 / 主编

解放军外国语学院英语博士文库

Learner Behaviour of Verbs: Colligation, Collocation and Chunk

濮建忠 著

学习者动词行为：
类联接、搭配及词块

河南大学出版社

Monographs by Ph.D Holders,
English Department, PLA UFL

H314
P819



郑州大学

04010227173R

程 工 / 主编

解放军外国语学院英语博士文库

-7

Learner Behaviour of Verbs: Colligation, Collocation and Chunk

濮建忠 著

江苏工业学院图书馆
藏书章

为:
类联接、搭配及词块

河南大学出版社

Monographs by Ph.D Holders,
English Department, PLA UFL



H314

P819

2004/9/15

图书在版编目(CIP)数据

学习者动词行为:类联结、搭配及词块/濮建忠著. —开封:河南大学出版社,2003.12(2004.7重印)
(解放军外国语学院英语博士文库/程工主编)
ISBN 7-81091-176-7

I. 学… II. 濮… III. 动词-研究-英文 IV. H042

中国版本图书馆 CIP 数据核字(2003)第 001936 号

书 名 学习者动词行为:类联结、搭配及词块
作 者 濮建忠

出 版 人	王刘纯	责任编辑	薛巧玲
责任校对	苏 婷	责任印制	苗 卉
装帧设计	张 胜·生生书房		

出 版	河南大学出版社		
	地址:河南省开封市明伦街 85 号	邮 编:	475001
	电话:0378-2864669(行管部)	0378-2825001(营销部)	
	网址:www.hupress.com	E-mail:	bangong@hupress.com
经 销	河南省新华书店		
排 版	河南大学出版社印务公司		
印 刷	河南第一新华印刷厂		
版 次	2003 年 12 月第 1 版	印 次	2004 年 7 月第 2 次印刷
开 本	650mm×960mm 1/16	印 张	18.75
字 数	255 千字	印 数	2001—4000 册

ISBN7-81091-176-7/H·116 定 价: 21.00 元

(本书如有印装质量问题,请与河南大学出版社营销部联系调换)

序 言

语料库语言学：渊源，理念，方法及其他

濮建忠博士要我为他的专著《学习者动词行为：类联接、搭配及词块》作序，顿使我诚惶诚恐。一则，就研究专题而言，中国大学生的英语动词使用，我很少涉猎；二则，就研究方法而言，作者的语料库研究技艺，我不过仅能望其项背而已。然而，作为同道者，我对作品的学术分量却深信不疑，对书中展示的研究方法推崇备至。那么，我就捉刀代笔，在此谈两个问题：一是作者为何化数年的艰辛劳动，用语料库方法研究中国学生的英语动词使用？二是该研究给我们揭示了什么？前者涉及语料库研究的思想渊源、语言学理念与研究方法等等，后者涉及语料库研究的主要环节、步骤和特征。

先说第一。濮建忠博士为进行此项研究，首先进行了专门的数据建设。在桂诗春、杨惠中两位先生的指导下，他承担了中国学生英语语料库 CLEC 建设的重要任务，对学生的作文语料科学抽样、编辑整理、标注赋码和统计分析。工程之浩繁、过程之漫长、劳作之艰辛，绝非寥寥数语可表。由此建成的 CLEC 理所当然获得了学界众口赞誉。濮建忠博士为 CLEC 付出了可贵的努力，也为他的此项研究构筑了坚实的数据基础。那么，有必要为此花费巨力，建设专门的语料库吗？语料库语言学到底是个什么样的学科，使

研究者如此不遗余力呢？

人们常说，语料库是语言研究的一个工具，语料库语言学因此也是一门关于工具的学科。此言不差。但这“工具”和“工具学”却反映了一种重要的学术思潮，代表了一种语言学理念，揭示了一种研究方法，形成了一种全新的研究事业。在哲学层面上，自古即有“理性主义”与“经验主义”之争。理性主义认为，人对客观世界的认识或知识来源于大脑内部；内省是一切知识的基础；真理靠感知而来。经验主义则认为，人的客观世界经验和活动是一切知识之源；真理由求证而知。反映到 20 世纪的语言学领域，这两种哲学思潮形成了两大语言学流派，一为乔姆斯基开创的“形式主义”，一为韩礼德代表的“功能主义”。乔氏不愧为语言思想的巨人，他视语言学为心理学研究的一部分，视语言运作为人之生物学属性之体现，创设了一套庞大的转换生成规则，将纷繁复杂的语言使用高度形式化，以诠释人类理解和使用语言的内在机制，十数年内模型几经修订，可谓志在必得。而韩氏则视语言学为社会学研究之一部，视语言为一种社会符号系统，将语言使用作为一种社会行为，十数年内由系统再功能，皇皇巨著，屡创经典。由此形成的两军对垒在研究方法上也泾渭分明。形式主义以内省数据为其基本依据，以理想本族语者的直觉判断证据是否合法；功能主义则承继弗思语言学的实证研究传统，以语言交际活动中业已使用的真实证据为主要依据。

那么，从思想渊源、语言学理念到研究方法，语料库语言学何属何从呢？我们不妨认为，语料库语言学隶属于经验主义的一派学术思潮，它注重人的客观世界经验，强调语言使用者的实践，因而与理性主义相对。在语言学流派上，语料库语言学认同于功能主义，与形式主义相对；它视语言为一种社会行为或社会现象，而非一种心理现象。视语言为一种社会现象，就意味着它的研究对象为外部语言，而非内部语言。外部语言是可观察的、可记录

的、可描述的和可分析的文本数据。虽然语料库数据也可被用来研究人们的心理词汇等心理语言学问题，但语料库语言学者并不刻意进行此类研究。在他看来，究其本质，内部语言是难以观察和描述的，而相比之下，大量真实的人类语言交际活动尚未得到有效、科学的记录和描述，这显得格外紧迫和重要。就研究方法而言，语料库语言学者开发了完整、系统的工作方法。首先要根据研究目的和内容，对研究对象科学地抽样、采集证据，按照一定的语言学原则编制具有代表性的语料库，并进行必要的标注或赋码。继之，要用检索工具从语料库自动提取数据，并进行统计学处理和分析。最后，对数据进行语言学解释和描述。

语料库语言学者不排斥直觉的作用。事实上，在研究过程中，人们都在不同程度上运用着个人的直觉。但语料库语言学者却不使用直觉数据或内省数据。在他看来，数据只能从自然交际活动中忠实记录，而不能生造；本族语者虽有能力判断句子的合格与否，但个人的语言直觉实在有限，当一个语言学家为自己假设的语言学模型自造证据时，会苦思冥想出某些在其他本族语者看来并不合法的句子。况且，人脑会玩弄花招，任何不同寻常的特殊语言现象会被格外注意，最为普通和一般的常见现象却会被视而不见。当自造数据不再合格和有效，由其支撑的理论模型还有多大解释力呢？因此，数据建设便成为重中之重。真实数据，语言交际活动中的自然数据，而非语言学家头脑中的数据，是语言描述和语言学模型构建的生命所在。因而，语言研究应基于真实证据，而非其他。诚如充满睿智和创造精神的辛克莱教授所针砭：一个植物学家不能靠观察塑料假花去研究植物学。然而，并非进到篮子里的都是菜。语料库语言学者依赖的是具有代表性的自然数据，他的数据样本应具有代表语言整体的属性，他从语料库中提取的数据应具有统计学意义上的显著性。由此，建库格外耗时耗力，数据取舍格外谨慎劳苦。

从另一方面说, 不要自然数据也可研究语言问题, 而且在某些方面或许较为省力。从孔夫子以来, 或者说从汉代董仲舒“罢黜百家, 独尊儒术”以来, “述而不作”可谓我们学界的主流思想。从书本到书本, 从理论到理论, 能将古人、贤人的经典之说读懂、述透, 已属不易, 何苦再另作一番呢? 所以, 我们看到, 学界引进理论之风日盛, 模型层出不穷; 诠释、思辨、博引者多, 处理、分析数据者寡。当然, 读书无疑极为重要, 尤其是读懂古今中外的经典之作。倘非如此, 先人创造的文明成果、大师们卓越的学术成就将无以为继。所以, 我们对破书万卷、学识渊博的长者和同仁充满了由衷的敬意。但是, 这些都不应取代基于真实语言交际证据的概括、归纳和抽象。

至此, 我们可以回答篇首提出的第一个问题了。濮建忠博士为研究中国学生动词用法花费巨力、巨时建设语料库, 实属必要, 是语料库语言学思想和理念的要求。业已证明, 桂诗春、杨惠中两位语言学家领衔建立的 CLEC 语料库是个具有良好代表性的语料库。本书作者基于该语料库进行的中国学生英语动词用法研究也是有效的研究。

关于第二个问题, 我们先领略一下作者采用的主要研究方法和步骤:

1. 建立学习者语料库并选择控制语料库, 进行对比研究;
2. 选择、确定欲研究的关键动词;
3. 使用索引软件检索出 CLEC 和控制语料库中关键动词所有实证, 并随机抽样;
4. 观察每一索引行, 标注类联接后重新排序, 获得类联接使用的有关统计数据;
5. 就主要的类联接进行搭配研究, 用半自动方法提取搭配词 (Collocates);
6. 基于数据, 对比研究学习者与本族语者在类联接和搭配层面

上的动词行为异同;

7. 分析产生差异的原因, 探讨对语言研究和语言教学的启示。

这些步骤揭示了语料库语言学研究的主要方法特征: 证据检索、抽样、统计等采用自动或半自动手段, 尽量避免人为干扰因素; 数据的解释、概括和归纳则要定量与定性研究相结合。量化的数据是基础, 定性分析则是语言学抽象, 依据一定的理论和观点进行。作者依据大量统计数据, 分析了动词诸用法“多用”、“少用”和“误用”等特征, 对中国学生的动词行为模式进行了抽象, 就词汇和语法的共选关系、母语的干扰因素、词块的性质与作用提出了真知灼见。

那么, 该书向我们展示了一个语料库研究的良好范例, 揭示了语料库语言学的主要理念和方法。作者的艰苦劳动结出了丰硕果实。我们为书的出版欣慰。语料库语言学呼唤更多、更好的同类研究问世!

卫乃兴

2003年6月

前 言

词汇 (Lexis) 和语法 (Grammar) 是语言学习除语音之外的两大要素。英语的教和学自然也回避不了英语词汇和语法。但是, 词汇和语法这两个语言中密不可分的方面却被许多语言学家和语言教师人为地进行了不恰当的分割。就语言理论而言, 在词汇和语法的处理上, 多数有重语法轻词汇的倾向, 而且两者之间的界限往往被认为是泾渭分明的, 两者也因此被不合理地割裂开来。就语言教学而言, 虽然语法翻译法早已不是主流, 但语法在课堂教学中仍有相当重要的位置却是不争的事实, 词汇在课堂教学中却远远没有得到应有的重视。更为关键的是, 语法和词汇在课堂教学中是相互独立而不是有机联系的, 两者间的密切关系被严重忽视。这种对词汇与语法之间关系的认识不足在一定程度上制约了许多语言理论的发展, 同时也在很大程度上限制了英语教学的改进和提高。

语料库语言学 (Corpus Linguistics) 自 20 世纪 80 年代以来的蓬勃发展为进一步揭示词汇和语法之间的关系创造了现实条件。当代语料库语言学把语言的实际使用放在了首要的位置, 以海量真实的语言数据为驱动, 采用人机相结合的方法, 充分利用计算机巨大的存贮容量, 高速、有效、准确的运算和信息处理能力以及语言研究者的概括、归纳、推理、直觉、创造等能力, 应用统计学的原理使定性与定量研究相结合, 揭示了语言中本身固有的却又通过其他方法无法提炼的概率信息, 在语言学上引起了一场

革命。语料库语言学的研究成果表明,语言及语言使用中的形式(Form)和意义(Meaning)密切相关。离开意义谈形式,或者离开形式谈意义都无法揭示语言和语言使用的本质。意义和形式之间虽然不是一一对应的关系,但一定的意义往往用一定的形式来表示,一定的形式也通常表示一定的意义,两者之间的紧密联系不容忽视。就词汇和语法的关系而言,两者是语言形式不可分割的两个方面,都是为表达意义服务的,而且句法(Syntax 即狭义上的语法)与词汇是共选(Co-selection)关系。换言之,当语言使用者用语言表达某一意义,传递某一信息时,他们往往想到某一或者有限的几种语言形式,在形式的构成过程中,既不是先确定句法结构再选择词汇,也不是先确定词汇再选择句法结构,通常的情况往往是句法结构和词汇共同选择,实现表达意义的目的。“形式和意义密切相关,句法与词汇共选”的思想与著名语言学家韩理德(M. A. K. Halliday)于20世纪60年代明确提出的“词汇语法”(Lexico-grammar)思想基本一致,语料库语言学的研究成果将词汇语法不断引向深入。

本研究正是利用语料库语言学的研究方法,以自行建立的学习者语料库(Learner Corpus)和几个国内外具有代表性的本族语者语料库(Native Speaker Corpora)的证据为基本依据,采纳词汇语法的主要思想和观点,从类联接(Colligation)和搭配(Collocation)两个层面对中国大学生使用英语动词的情况进行系统深入的观察、描述、分析,以期概括和归纳出中国学习者英语动词使用的行为模式,为英语二语习得(Second Language Acquisition)尤其是学习者中介语(Interlanguage)的研究,以及外语教学提供借鉴。利用大规模学习者语料库和本族语者语料库进行定性 with 定量相结合的学习者中介语研究在国际上尚属开始阶段,在国内更是寥寥无几。以动词为切入点,从动词使用所涉及的两个互为联系的关键层面——类联接和搭配,来研究学习者的动词行为还很少见。

更值得一提的是,本研究不仅探讨了学习者在动词使用过程中在类联接和搭配上的错误(Misuse)或不当(Inappropriacy),而且还探讨了在类联接上学习者过度(Overuse)和偏少(Underuse)使用的情况,从而更加全面和精确地反映了学习者中介语不同于本族语的特点。

本研究发现,中国学习者在英语常用动词的使用中,出现在类联接层面上的错误还相当普遍,出现在搭配层面上的不当似乎更为严重,出现在前一个层面上的过度使用和偏少使用也极具普遍性。经过对上述结果的分析和探讨,作者发现,中国学习者在英语常用动词的使用中在类联接和搭配两个层面上所出现的问题可以从学习者所掌握的相对完备的母语(汉语)知识及其相对部分的、不完备的目的语(英语)知识对其中介语的影响上得到较好的解释,即:学习者的动词行为可以通过语际迁移(Inter-linguistic Transfer)和语内迁移(Intra-linguistic Transfer)来进行解释。进一步的分析和探究还显示,问题的关键并不在于学习者对某一动词的意义、或者它的类联接、或者它的搭配等的某一个或几个方面掌握不好,而是在于学习者普遍没有把意义和形式有机结合起来,没有把动词的意义和类联接、搭配等融为一体进行学习。相反,许多学习者在学习英语动词(其他词类也如此)时,常常把学习它们在汉语中所谓的“对等词”(Equivalents)或同义词作为重要目的,甚至是终极目的。这种忽略词的最直接的语内语境(Immediate Linguistic Context 其中两个最重要的方面就是类联接和搭配)、孤立学习目的语单词在母语中的对等词或同义词的学习方法不可能保证学习者掌握词的确切意义,更谈不上掌握其在目的语中的地道用法。要把意义、类联接和搭配融为一体,必须引入一个新的概念。本书作者在前人相关研究的基础上,提出了“词块”(Chunks)这一概念,并对它进行了工作定义。词块这一概念体现的正是词汇语法的基本精神和观点,它使类联接和搭配有机

地结合在一起, 并对中国学习者的英语动词行为做出了更全面、更准确、更具说服力的解释。

本研究的一条重要结论是: 中国学习者在英语动词使用上所表现的不足, 究其根本原因还在于学习者对英语中能体现动词用法的典型词块掌握不够。作者认为, 本研究的结论不仅适用于动词, 还适用于其他词类; 不仅适用于英语, 也适用于其他语言作为第二语言或外语的习得。为此, 本研究对语言教学和学习提出了一个建议, 即: 语言教学和学习的一个主要焦点应放在常用词及其构成的常用、典型的词块上。

为便于读者参阅此书, 下面就各章的主要内容做一简要的介绍。

本书第一章论述语料库语言学的主要特点(即: 语料库文本的真实性、语料库数据的可计算性、语料库文本中的概率性), 描述采用语料库语言学方法之前二语习得研究领域各类分析方法的发展(包括对比分析、错误分析、中介语应用分析、语篇分析法), 指出各类方法不可避免缺陷, 阐明学习者语料库在二语习得研究中的重要位置和价值。学习者语料库的产生孕育了一种新的分析方法, 即“对比中介语分析”。该方法的应用使得系统、详尽、定性与定量相结合地研究学习者的语言行为成为可能。

第二章首先描述动词在传统语法及两本重要的教学语法(*A Comprehensive Grammar of the English Language*, *Collins COBUILD English Grammar*)中的处理方式, 指出在动词处理方式上的不同反映的是对词汇和语法之间的关系的不同观点和态度。接着阐述词汇与语法之间的关系, 指出两者不能割裂, 而应结合在一起, 由此引出了词汇语法的思想, 并且认为对动词的更好描述和处理应该从词汇语法的观点入手, 而要如此处理动词则要考虑动词使用时的最直接的语内语境(其最重要的两个方面就是类联接和搭配)。

第三章以阐述研究方法为主要目的。首先解释本研究中所使用的语料库（包括学习者语料库、用于对比分析的参照控制语料库）及其主要特征，然后描述被研究动词的选择过程（该过程需要考虑诸多因素，如动词意义、用法、常用性、学习者对它们的掌握情况等等，目的是抽取一个具有一定代表性且能充分反映学习者对其掌握情况的样本），最后对动词使用的两个重要方面——类联接和搭配，进行必要的解释。

第四章介绍和初步探讨对选择出的英语常用动词样本的研究结果，描述学习者动词行为的显著特征，并从语际迁移和语内迁移两个方面探讨产生这些特征的根源。结果和讨论分三部分展开：第一部分只对单个动词的行为进行描述；第二部分对同一语义域且意义相近的动词行为进行了对比；第三部分对同一动词不同形式的行为进行了对比。研究得出的初步结论是，从学习者的动词类联接和搭配行为来看，这一层次的中国英语学习者还在很大程度上受到母语的影响。不论是类联接上的错误使用、过度和偏少使用，还是搭配词不当或非典型的选择，不管是近义词相同或不同的使用，还是动词的不同形式的近似使用，都与母语的影响有很大的关联。此外，还有一些受目的语本身影响的证据。但是，研究也发现了学习者正确和贴切使用动词的许多证据。这类正确、贴切的用法掌握得越多、越牢靠，语际迁移和语内迁移发生的可能性就越小。

第五章在前一章的基础上对所得到的结果做了进一步的分析、探索和思考，发现问题的关键并不在于学习者对某一动词的意义、或者它的类联接、或者它的搭配等的某一个方面或某几个方面掌握不好，而是在于学习者普遍没有把意义和形式有机结合起来，没有把动词的意义和类联接、搭配等融为一体进行学习。为了使意义、类联接和搭配融为一体，笔者在前人相关研究的基础上，提出了“词块”（Chunks）的概念。本章首先介绍词块概念的来由，

并对它进行工作定义, 然后通过这一概念对学习者的动词行为做出进一步的解释, 最后论述本族语者典型词块如何为学习者所习得。笔者发现, 学习者与本族语者在动词类联接和搭配行为上的主要差异可以归结为他们所使用的典型词块之间的差异。学习者所使用的词块在质和量上不同于本族语者的词块的主要原因, 一方面是由于许多学习者在相关汉语的词块基础上产出英语词块, 另一方面是因为当他们对目的语的部分知识不足以告知他们正确的英语词块时, 做了不正确的类比 (Incorrect Analogies) 或过度概括 (Overgeneralizations)。因此, 本研究指出, 学习者动词行为有别于本族语者动词行为的一个根本原因是学习者对英语中能体现动词用法的典型词块掌握不够。

本书最后一章, 首先对前面的内容做了一个总结, 然后探讨了本研究对语言研究和语言教学的启示, 结尾提出了一些进一步研究学习者中介语的思路和建议。本研究对语言研究的主要启示有: 1) 本研究结果支持辛克莱提出的关于成语原则 (Idiom Principle) 的论断; 2) 本研究支持词汇与句法相互依赖、词汇与句法共选的观点; 3) 本研究采用了词汇语法的基本观点来综合研究动词的类联接和搭配行为, 研究结果反过来也证实了词汇语法观点的正确性。本研究对语言教学的主要启示有: 1) 常用词汇在语言教学中应该得到高度重视; 2) 本研究指出了词汇教学的两个关键方面——类联接和搭配; 3) 本研究提出, 语言教学和学习的一个主要焦点应放在常用词及其构成的常用、典型的词块上。

本书的写作得到了本人导师、上海交通大学杨惠中教授的悉心指导, 也得到广州外语外贸大学桂诗春教授的鼓励和指点, 同时还受益于和卫乃兴、李文中博士的思想交流。父母和妻子无私的爱以及不尽的支持给了作者完成此书的勇气和信心。这本小书能跟大家见面还要归功于解放军外国语学院程工教授、潘永樑教授、李绍山教授、严辰松教授、李经伟教授等一大批德高望重的

老师，他们多年来对作者的勉励和关怀使作者如沐春风，并给了作者在语言教学与研究领域孜孜以求的强大动力。另外，本书的出版得到了河南大学出版社的大力支持，书中也凝聚了编辑们的辛勤汗水。在此作者一并对他们深表敬意和感谢。

限于作者本人的水平，错误和纰漏在所难免，在研究方法和手段上也可能存在一些欠缺或不足，恳请专家同行批评指正。

濮建忠

2003年7月

Contents

Chapter One	Corpus Linguistics and SLA Research.....	(1)
1.1	The Characteristic Features of the Corpus Approach	(2)
1.2	Types of Analysis in SLA Research	(18)
1.3	Learner Corpora: the Meeting-point of SLA and Corpus Linguistics.....	(33)
1.4	Summary	(46)
Chapter Two	Approaching Verbs	(47)
2.1	Grammar and Verbs	(48)
2.2	Grammar and Lexis	(57)
2.3	Verbs and Its Immediate Linguistic Context.....	(73)
2.4	A Comparison of Chinese Learners' Behaviour of Verbs with That of the Native Speakers.....	(79)
2.5	Summary	(85)
Chapter Three	Research Methodology.....	(87)
3.1	Corpus Data	(88)
3.2	The Verbs to Be Investigated	(97)
3.3	Two Aspects of the Use of Verbs	(118)
3.4	Summary	(130)

Chapter Four Research Results and Discussions	(132)
4.1 The Uses of Individual Verbs.....	(133)
4.2 Comparisons Between Uses of Verbs	(167)
4.3 The Uses of the Different Forms of Verbs	(178)
4.4 Some General Conclusions	(188)
4.5 Summary.....	(189)
Chapter Five Further Discussions	(191)
5.1 The Notion of Chunk	(191)
5.2 Chunks and the Learners' Uses of Verbs	(194)
5.3 The Learning of the Native Speakers' Typical Chunks	(202)
5.4 Summary.....	(212)
Chapter Six Conclusions and Implications	(213)
6.1 Summary and Conclusions	(213)
6.2 Implications	(217)
6.3 Directions for Further Research.....	(232)
Appendices	(236)
References	(275)