

# 搭配研究 50 年：概念的演变与方法的发展

卫乃兴

(上海交通大学语言文字工程研究所，上海 200030)

**摘要：**在过去半个世纪里，搭配 (collocation) 经历了概念的演变和方法的发展。主要的理论体系包括弗思学派的概念和研究方法，米切尔等人的综合法，韩礼德和哈桑的篇章衔接概念，博林杰等人的惯例化搭配研究，以及语料库研究方法。本文讨论各研究体系的界定特点及其差异，并概括概念演变的脉络与方法发展的趋势。

**关键词：**搭配；共现；类联接；搭配衔接

中图分类号：H313 文献标识码：A 文章编号：1002-722X (2003) 02-0011-05

## Fifty Years of Collocational Study : Conceptual Changes and Methodological Developments

WEI Nai xing

(Natural Language Processing Institute, Shanghai Jiaotong University, Shanghai, 200030, China)

**Abstract :** This paper presents a review of the major approaches to the study of collocation in the past fifty years, including the Firthian lexical approach, Mitchell's integrated approach, Halliday and Hasan's concept of collocational cohesion, Bolinger's concept of institutionalization, and the corpus based approach. Major defining features of each approach are discussed, and the general trend for conceptual changes and methodological developments summarized.

**Key words :** collocation ; co occurrence ; colligation ; collocational cohesion

### 1. 引言

搭配 (collocation) 无疑是语言学领域最为重要的概念与研究内容之一。在其问世后的半个世纪里，界定体系与研究方法不断演变和发展：概念繁多，异质性大，研究方法迥异。主要体系与方法有：弗思 (J. R. Firth) 于 20 世纪 50 年代提出的搭配概念、新弗思学派于 60 年代提出的词语学研究体系、米切尔 (T. F. Mitchell) 等人于 70 年代主张的综合法、韩礼德和哈桑 (1976) 的篇章衔接概念、博林杰 (Bolinger, 1976) 等的惯例化搭配研究，和 80 年代兴起的语料库语言学方法。

系统梳理上述主要概念、分析各研究体系的特点及其存在问题，并概括概念演变的脉络与方法发展的趋势，是语言学研究的重要课题。

### 2. 弗思学派的搭配概念与研究方法

#### 2.1 弗思提出的搭配概念：**词项结伴、相互**

#### 期待、类联接

弗思被公认为搭配研究之父。(Mitchell, 1975, 134) 他提出的重要概念之一就是“结伴”说——搭配即词项的习惯性结伴使用。在其语言学论著中，我们可以多次读到“由词之结伴可知其词”的著名论断。(Firth, 1957, 12) 弗思认为，结伴的词项体现了一种意义方式；词项有多种意义方式，可分别在正字法层面、语音层面、语法层面和搭配层面进行研究。而在搭配层面，词项的意义之一就是与别的具体词项结伴使用。比如，“night”的意义之一就是与“dark”结伴。(ibid: 182) 也就是说，搭配是关于词项组合的研究。然而，搭配不仅仅是词的简单并置，它体现了词与词的相互期待 (mutual expectancy)。如“dark”与“night”相互期待。高于搭配层面的抽象是类联接 (colligation)，即语法范畴间的组合关系。搭配研究的是具体的词与词的

收稿日期：2002-12-26

作者简介：卫乃兴 (1955—)，男，河南济源人，上海交通大学语言文字工程研究所教授，博士，主要研究方向为语料库语言学和应用语言学。

(C)1994-2022 China Academic Journal Electronic Publishing House. All rights reserved. <http://www.cnki.net>

组合，而类联接则仅涉及抽象的语法概念与范畴。

弗思的搭配概念无疑具有丰富内涵。他所说的词项结伴，即今天语言学者谈的词项共现，是词汇行为的重要物理现象。将搭配界定为对词项组合关系的研究，是十分重要的学术立场：组合不同于聚合，后者涉及词义间的联想关系，涉及词项搭配行为的可能性；前者则涉及其现实性与典型性。搭配伙伴的相互预见性也是研究的重要内容。一方面，在语言使用中，习惯性搭配的词项在文本中相互预见：一个伙伴一经出现，另一个伙伴极有可能也在文本的某个地方出现。另一方面，在研究实践中，采取某种手段确定了词项间存在着相互预见，便可确定它们为搭配。然而，弗思并没有对如何研究词项的相互预见性进行探讨。类联接是重要的句法概念。但是，类联接与搭配的关系如何？究其本质，作为句法层面上语言组合关系的抽象，类联接的成分也相互预见，那么类联接与搭配也必然相互影响和制约。然而，弗思并未对此系统地探讨。他似乎认为，搭配研究可以独立于句法关系进行。(Firth, 1957) 在研究方法上，弗思主张使用验证数据，即业已使用过的真实证据，反对生造例证，并认为“用机器分析文本是可能的”。(Firth, 1957)

## 2.2 新弗思学派的词语学体系：词项、搭配、词语集

新弗思学派 (Neo Firthians) 的代表人物是韩礼德和辛克莱 (J. Sinclair)，另有麦金托什 (A. McIntosh)、斯特雷文斯 (M. Strevens) 等人。麦金托什 (1967) 将搭配范围 (collocational range) 区别于语法型式。后者揭示句子结构的合格性，前者则揭示词项兼容其他词项的搭配力。任何词项都只有有限的搭配力：形容词 “molten” 与 “metal、iron、lava” 等词兼容组合，但不与 “feather、postage、pig” 等词兼容组合。搭配范围与语法型式同样重要，其研究需用一种词语学的方法。

韩礼德 (1966) 和辛克莱 (1966) 是搭配研究进程中两篇具有划时代意义的文章。他们提出了词语学 (lexis) 研究中三个密切关联的概念，词项 (lexical item)、搭配和词语集 (lexical set)；三者相互界定、相互依存。韩礼德 (1976: 75) 将搭配界定为“体现词项在某种显著的临近范围内组合关系的线性共现”。该定义明确无疑地将搭配界定为组合关系，将“线性共现”视为搭配最重要或者惟一的界定标准。这无疑是对弗思“结伴”说的进一步阐释。据此，句子

中的“strong argument” 是搭配，但是，

(2) I wasn't altogether convinced by his argument. He had some strong points, but they could all be met.

中的“strong” 和 “argument” 也是搭配，因为二者存在一定程度上的线性共现。

词项是个抽象的语言单位，其认定取决于相关词形搭配型式的异同。据此，“strong”，“strongly”，“strength” 和 “strengthen” 等可被视为同一词项 “strong”的不同词形，因为它们的搭配型式一样，(Halliday, 1976: 74) 如下面 (3) 至 (6) 句所示：

(3) He argues strongly.

(4) He put forward a strong argument for it.

(5) I don't deny the strength of his argument.

(6) His argument was strengthened by other factors.

词语集是具有相同搭配型式的有关词项的集合。(Halliday, 1966: 153) “strong” 和 “powerful” 都可与 “argument” 搭配，因而都属于由 “argument” 限定的词语集。但二者并不总属于同一词语集。在由 “tea” 限定的词语集里，“strong” 是个成员，“powerful” 则不是。词语集体现词项的聚合关系，属于语言的系统，搭配则体现词项的组合关系，属于结构。

韩礼德 (1966) 和辛克莱 (1966) 认为，搭配研究必须基于语料库证据。为此，他们提出了一整套概念和方法，如节点词、跨距、搭配词，从语料库提取搭配例证。他们发展了弗思关于搭配伙伴相互预见的论说，认为节点词与搭配词具有相互吸引力 (mutual attraction)，并用一套统计手段测量之，从而确定搭配关系。(卫乃兴, 2002a) 这套方法给搭配研究带来了巨大变化；这一点将在本文第 6 节中详细讨论。

显而易见，新弗思学派的词语学方法是对弗思概念的继承与发展。弗思关于搭配的组合界定特点得以继承：结伴说被阐述为“线性共现”；“相互预见” 观点发展为可计算的“相互吸引力”。新弗思学派主张的基于语料库的研究方法确有独创之处：它是弗思“使用验证数据” 主张的具体实现，使搭配研究建立在了真实语言使用的路上，从而保证了效度。另一方面，该体系也有弱点。其一是概念界定中的循环论证：词项依照词形的搭配型式界定；反过来，搭配要依照词项界定。其二涉及搭配的界定标准。对一般语言学者来说，线性共现可能只是“物理”的界定标准，搭配伙伴间的句法与语

义关系似乎更应考虑。人们会认为，上述例(2)中的“strong”和“argument”并非搭配，倒是“strong points”应被视为搭配，因为后者具有句法互限关系。然而新弗思学派认为，将句法关系纳入界定标准非但不能增加体系的解释力，还会削弱之。另外，在界定体系中避而不谈语义关系自然会受到语义学家质疑。莱昂斯(1966: 288—297)批评弗思学派的理论‘不足以研究搭配’，“将意义仅仅界定为语境关系的组合体令人困惑”。语义学家趋向于用选择限制机制解释词项组合。(Leech, 1981: 138—139; Gruse, 1989: 104—106)照此机制，“love”要求其主语具有[Animate]特征，“divorce”要求其主语为[Hu-man]等。然而，该机制却解释不了为何“flawless argument”、“spotless reputation”等是英语的合适搭配，而“flawless reputation”、“\*spotless argument”等则不是。说到底，搭配具有因循性或无规则性的特点。新弗思学派在界定体系中摈弃句法和语义标准似乎有其道理。究其本质，该方法是他们词语学体系的一部分，不直接涉及句法与语义。

### 3. 米切尔等人的综合法：词根、句法、意义

格林鲍姆(Greenbaum, 1970)等认为，新弗思学派的词语学方法是一种面向词项的方法，而搭配研究需要的则是一种综合方法，因此提出将语法结构和句型纳入搭配研究。米切尔(1971, 1975)则系统地论述了搭配研究的综合法，将词根、句法和意义纳入了研究。

首先，米切尔认为，搭配存在于词根之间，而非词之间。词根是一列词形所含的共同部分，或最高公因式HCF，用√表示。据此，√drink是“drinks”，“drink”，“drinking”等词形的共同词根；√heavy则是“heavy”和“heavily”的共同词根。下列(7)至(10)句都存在√drink与√heavy的搭配，用√drink√heavy表示。(Mitchell, 1975: 118)

- (7) He is a heavy drinker .
- (8) He is drinking pretty heavily .
- (9) He drinks heavily .
- (10) He is putting in some heavy drinking .

再者，搭配受类联接限制。米切尔(1975: 83)认为，类联接即搭配类(collocational class)，涉及使用词类给搭配冠名；搭配则是类联接的具体成员。他对词组动词“tore up”与其后接宾语名词所构成的类联接概括如下：

- (11) He tore up the paper . → Colligation : V +

### Particle + (travelable) N

由上述两个不同类联接可知，米切尔的类联接概念已有别于弗思，已将词类、句法、意义等因素综合一体。

综合法易为学术界接受。考伊(Cowie, 1978)、谢尔默(Kjellmer, 1984)等，都用了类似的界定方法。然而，如果说新弗思学派将句法排除在界定特征之外有其弱点，那么，将搭配严格限定于句法结构内，也难免对许多搭配现象的研究带来问题，如下例中黑体词组成的非连续性搭配：

- (13) He didn't want the **job** . I don't think he even **applied** . (Mitchell, 1975: 120)

此外，“搭配存在于词根之间”未免失之于过度概括。具有共同词根的诸词形，搭配行为并不都相同。究其本质，米切尔的综合法是其一般语言理论体系的一部分，侧重于语言的定性分析，不太易于定量研究。

### 4. 韩礼德和哈桑的篇章衔接概念：成对词、搭配衔接

韩礼德和哈桑(1976: 287)认为，许多成对成串的词项常出现于文本的同一语境，由此产生篇章衔接作用，可称为搭配衔接；这种衔接与其说产生于词项间的语义关系，不如说产生于相互搭配。如laugh … joke、cradle … flame … flicker、sky … sunshine … cloud … rain等。

由此可知，韩礼德和哈桑一方面坚持认为，搭配可以不考虑句法与语义关系，而仅靠“共现”确定。另一方面，他们发展或改变了弗思和新弗思学派的概念。似乎只要这些成对词出现于共同语境，就会产生篇章衔接作用，因而也就是所谓搭配。然而，我们不难发现，上述“成对词”有着巨大的异质性，多指客观世界里相互联系的事物，内在关系复杂。如果将它们视为搭配，就会混淆搭配的概念。就研究方法而言，新弗思学派主张的实证研究，一般都要界定共现词语的语境，用统计手段测量共现数据，以确定搭配的典型性。但，这里提出的成对词一般都离散地分布在文本中，使语境界定和实证统计难以进行。至少就目前的语言学理论水平和研究手段而言，这种篇章衔接概念的搭配难以付诸研究。

但是，词语搭配的篇章衔接作用毋庸置疑，韩礼德和哈桑的观点对挖掘词语搭配的丰富内涵确具有有益的启示。

### 5. 博林杰等人的惯例化搭配研究：预制品、定型表达、程式语

Particle + (destructible) N

博林杰 (1976) 视搭配为一种语言预制件或定型表达。即搭配像成语一样, 是一种语言使用惯例或规约。如, 人们使用 “the livelong day”, 却不用 “<sup>a</sup>the livelong morning” 等。这就是词语搭配的惯例化属性或因循性, 很难解释。博林杰是从语言习得和记忆储存的角度探讨搭配的。他认为, 词素可被储存, 也可能不被储存。但是, 搭配却是实实在在被储存的。在语言学习和使用中, 搭配被学习、记忆并整体提取使用, 形式与意义一体, 一切都同时发生。(ibid: 2) 他强烈批评那种赋予人类无限的语言生成能力但只有有限记忆能力的语言学理论: “语言并不期待我们从木材、钉子和蓝图开始去建造一切, 而是给我们提供了极其巨大数量的预制件。”(ibid: 9) 该批评无疑指向了居主流地位的转换生成语法理论。主流理论认为, 人掌握了一套有限的语法规则, 就可理解和生成无限量的句子。

波利和西德 (Pawley & Syder, 1983) 将惯例化搭配称为词语化句干 (Lexicalized Sentence Stem) 或形式—意义配对语。纳廷杰 (Nuttinger, 1988) 将“具有语用功能的词语搭配”界定为词语片语。这些实质上是些高度惯例化的搭配, 长度不等, 结构繁简不同, 是预先装配好的词组。而语言使用的过程就是将这些搭配缝制成语话的过程。考伊 (1988) 则称这类搭配为程式语 (formulae) 和复合语 (composites) 等。

关于惯例化搭配的研究揭示了词语搭配因循性的界定特点。可见, 搭配在一定程度上有其因循使用或约定俗成的性质。这种因循性也正是语言使用地道性的基础所在。那么, 任何认真的词语搭配研究体系都应考虑这一属性。

## 6. 语料库语言学的兴起: 自动检索、数据驱动、词语行为

琼斯和辛克莱 (Jones & Sinclair, 1974) 是第一个基于语料库的词语搭配研究报告。该研究确立了一系列重要原则和基本方法, 包括跨距界定、统计方法、搭配词与节点词相互吸引力的测量, 显著搭配的确定等等。(卫乃兴, 2002a) 20世纪80年代至今, 基于语料库或语料库驱动的词语搭配研究广泛展开。(Sinclair, 1987, 1991; Renouf & Sinclair, 1991; Kjellmer, 1984; Kennedy, 1990) 大量自动检索软件、自动统计软件和搭配程序被开发并应用于词语行为研究。其中, 辛克莱领导的COBUILD项目最为引人注目。这一时期的重要研究发现包括:

1) 一般文本中高频词的非词语化趋向 (delexicalization)

- 2) 语义韵现象。(详见卫乃兴, 2002b)
- 3) 由直觉认定的词的“核心意义”一般都不是语料库证据显示的高频意义。
- 4) 语言选择和文本组织主要基于成语原则。
- 5) 语法机制很难解释、分析文本构成。

简言之, 语料库语言学的兴起, 使搭配研究发生了巨大变化。首先, 研究效度提高。研究者可便捷地检索百万计乃至千万计词容量的连续文本, 概括关键词的搭配行为, 而不必求助于个人直觉, 真正实现了“使用验证数据”的原则。第二, 量化研究的比重增大, 从数据提取到数据处理等, 都有一套定量分析方法, 许多抽象的概念, 如“相互期待”、“搭配力”等都具有了可观察、可操作的定义。第三, 自动化程度愈来愈高, 研究过程中的人为干扰因素大大降低。最重要的是, 研究深度增加, 结果发人深省。很多项目已不仅限于词项组合行为研究, 已涉及文本的性质、语言选择、词语和句法、意义的关系等诸多重要语言学课题。

## 7. 结论

本文讨论了半个世纪来搭配研究的主要方法与概念。一方面, 这些方法和概念异质性较大: 在不同的研究体系里, 搭配可能意味着不同的内容; 每种方法可能只侧重于搭配的一个方面。在弗思学派的体系里, 搭配是词语的结伴或共现, 验证数据是基本依据, 统计测量是重要手段。在米切尔等人的综合法里, 搭配是词根与词根的组合关系体现, 句法限制等是必要的界定标准, 研究基于直觉。在韩礼德和哈桑的篇章衔接概念里, 搭配是产生衔接力的成对词语。博林杰等人的研究则侧重于搭配的因循性。另一方面, 词语搭配的概念演变脉络与方法发展趋向也比较明显。从弗思的结伴说, 到新弗思学派的线性共现界定标准, 到综合法主张的类联接框架内的词项组合行为研究; 从弗思的“意义方式”到“非词语化”趋向和文本研究; 从弗思的“相互期待”说到“相互吸引力”的计算, 概念愈益明晰、具体。传统的研究多基于直觉, 语料库研究则牢牢基于真实证据。词语搭配研究已延伸至或触及许多重要语言学课题。

## 参考文献:

- [1] 卫乃兴. 专业性搭配初探——语料库语言学方法 [J]. 解放军外国语学院学报, 2001, (4): 19—23.
- [2] 卫乃兴. 基于语料库与语料库驱动的词语搭配研究 [J]. 当代语言学, 2002, (2): 101—114.
- [3] 卫乃兴. 语义韵研究的一般方法 [J]. 外语教学与研究, 2002, (2): 19—23.

- 2002, (4): 300—307.
- [4] Bolinger, D. Meaning and Memory [J]. *Forum Linguisticum* 1.1 (1976): 1—14.
- [5] Cowie, A. P. The place of illustrative material and collocations in the design of a learner's dictionary [A]. In *In Honour of A. S. Hornby* [C]. Ed. Strevens. Oxford: Oxford University Press, 1978: 127—139.
- [6] Cowie, A. Stable and creative aspects of vocabulary use [A]. In *Vocabulary and Language Teaching* [C]. Eds. Carter and McCarthy. London: Longman, 1988: 126—39.
- [7] Cruse, D. A. *Lexical Semantics* [M]. Cambridge: Cambridge University Press, 1989.
- [8] Firth, J. R. *Papers in Linguistics 1934—1951* [M]. London: Oxford University Press, 1957.
- [9] Greenbaum, S. *Verb intensifier Collocations in English: An Experimental Approach* [M]. The Hague: Mouton, 1970.
- [10] Halliday, M. A. K. Lexis as a linguistic level [A]. In *In Memory of J. R. Firth* [C]. Eds. Bazell, Catford, Halliday & Robins. London: Longman, 1966: 148—162.
- [11] Halliday, M. A. K. Lexical relations [A]. In *System and Function in Language* [C]. Ed. Kress. Oxford: Oxford University Press, 1976: 73—83.
- [12] Halliday, M. A. K. & Hasan, R. *Cohesion in English* [M]. London: Longman, 1976.
- [13] Jones, S. & Sinclair, J. English lexical collocations: a study in computational Linguistics [J]. *Cahiers de Lexicologie* 23. 2 (1974): 15—61.
- [14] Kennedy, G. D. Collocations: Where grammar and vocabulary teaching meet [A]. In *Language Teaching Methodology for the Nineties* [C]. Ed. Anivan. Singapore: RELC, 1990: 215—229.
- [15] Kjellmer, G. Some thoughts on collocational distinctiveness [A]. In *Corpus Linguistics: Recent Developments in the Use of Computer Corpora in English Language Research* [C]. Eds. Aarts & Meijis. Costerus: N. S., 1984: 163—171.
- [16] Leech, G. *Semantics* [M]. Harmondsworth: Penguin Books Ltd, 1981.
- [17] Lyons, J. Firth's theory of meaning [A]. In *In Memory of J. R. Firth* [C]. Eds. Bazell, Catford, Halliday & Robins. London: Longman, 1966: 288—302.
- [18] McIntosh, A. Patterns and ranges [A]. In *Patterns of Language: Papers in General, Descriptive and Applied Linguistics* [C]. Eds. McIntosh and Halliday. Bloomington and London: Indiana University Press, 1967: 183—199.
- [19] Mitchell, T. F. Linguistic 'goings on': Collocations and other lexical matters arising on the syntagmatic linguistic record [J]. *Archivum Linguisticum* 2 (1971): 35—69.
- [20] Mitchell, T. F. *Principles of Firthian Linguistics* [M]. London: Longman, 1975.
- [21] Nattinger, J. R. Some current trends in vocabulary teaching [A]. In *Vocabulary and Language Teaching* [C]. Eds. McCarthy & Carter. London: Longman, 1988: 63—82.
- [22] Pawley, A. & Syder. Two puzzles for linguistic theory: Nativelike selection and nativelike fluency [A]. In *Language and Communication* [C]. Eds. Richard & Schmidt. New York: Longman, 1983: 191—225.
- [23] Renouf, A. and J. Sinclair. Collocational frameworks in English [A]. In *English Corpus Linguistics: Studies in Honour of Jan Svartvik* [C]. Eds. Karin & Bengt London: Longman, 1991: 128—143.
- [24] Sinclair, J. Beginning the study of lexis [A]. In *In Memory of J. R. Firth* [C]. Ed. Bazell, Catford, Halliday & Robins. London: Longman, 1966: 410—430.
- [25] Sinclair, J. Collocation: A progress report [A]. In *Language Topics: Essays in Honour of Michael Halliday* [C]. Eds. Ross & Threadgold. Amsterdam: Benjamins, 1987: 319—331.
- [26] Sinclair, J. *Corpus, Concordance, Collocation* [M]. Oxford: Oxford University Press, 1991.

(责任编辑 严辰松)