

## 《现代俄语详解组合词典》及其对汉语语义词典的启示

李 侠

(黑龙江大学俄罗斯语言文学与文化研究中心, 哈尔滨 150080)

**提 要:**《现代俄语详解组合词典》是莫斯科语义学派“意思⇔文本”模式理论的词库。《现代俄语详解组合词典》词条部分涉及的理论有语义元语言理论、支配模式理论和词汇函数理论。本文将《现代俄语详解组合词典》和汉语主要的语义词典相对照,分析《现代俄语详解组合词典》在词汇单位划分、释义、配价结构(支配模式)设置等方面为汉语语义词典提供的可借鉴之处,但在语义角色标注、语义搭配限制描写等方面仍有需要完善和补充之处。

**关键词:** 语义词典; 配价结构; 释义; 语义角色

**中图分类号:** H356

**文献标识码:** A

### 1 引言

《现代俄语详解组合词典》(Современный русский комбинаторный словарь 以下简称《组合词典》)出版于1984年,是莫斯科语义学派“意思⇔文本”模式理论(теория лингвистических моделей «Смысл⇔Текст»)的词库,它是理论语言学词典化的典型代表。《组合词典》词条部分分为释义区、支配模式区和词汇函数区,分别涉及的理论有语义元语言理论、支配模式理论和词汇函数理论,每个部分都有丰富的语言学理论做支撑。虽然《组合词典》出版至今已几十年,但其编写原则和结构设置对国内语义词典的编写仍具有借鉴意义。

### 2 《组合词典》对汉语语义词典的借鉴

对于动词配价词典,词汇单位如何设置、标明什么样的信息十分重要。范晓指出,理想的“动词配价词典”实际上是“某种语言生成句子的词典”,“这种配价词典应对某种语言里的动词配价组成的动核结构作出说明,不仅要给出某个动词(多义动词要分解到每个义项)的‘价类’,而且还要对动核的性质(动作、性状、变化、关系等等)和所联系着的动元(配价成分)的‘角色’(施事、受事、与事、系事、起事、止事……)以及某些常联系着的状元(非配价成分)作出准确的诠释;在此基础上,对该动词的形式特征以及它组配成的动核结构所生成的句模和各种可能的句型或句法格式进行详尽的描写。甚至,如果可能,还可注明该动词使用的条件和所组成的句型、句模组合体在语用上的选择性”。(范晓 1996:7)《组合词典》可以说是部分地实现了这一理想。

《组合词典》至少在以下几个方面可以为汉语语义词典提供借鉴:

#### 2.1 以义项为单位,进行自然语义元语言释义

语言精细化研究要求对词义作出更加细致的分辨,反映在词典中就是描写单位由“词”

转变为“词义”，即通常所说的“义项”。莫斯科语义学派常把多义词的一个义项叫做词汇语义单位（变体）（лексико-семантический вариант）（А.И. Смирницкий 1954； Л.А.Новиков 1982； Л.М. Васильев 1990）（本文简称词汇单位）多义词就是一个语义变体集。“以义项为描写单位描绘词的语义特征，建构语言知识库，正受到人们的高度重视”。（苏新春等 2005：162）《组合词典》把义项而不是词作为描写单位。

在给词汇单位（义项）进行释义时，《组合词典》采用自然语义元语言，将词汇语义分解成具有典型特征的语义元素，使语义的分解和表述更能够反映人们言语活动的心理表征和自然语感。元语言由语义单子和过渡语义元素构成，不仅使解释明确，而且简化了释文，更符合词典的编撰原则。例如：《组合词典》中 учиться “学习”的释义为：X учится Y-у(y Z-a/ по Z-у(в W-e)) = X 有意识地付出努力获取(Z 在 W 传授的)Y。’（Мельчук 1984： 917）该释文不仅提供了词汇单位基本的句法框架，释文中的变项还为配价结构的构建提供了依据和指定的信息。

目前汉语语义词典中往往选择性地对词汇释义进行，没有以义项为单位进行精细释义。同时没有对释义给予应有的重视，一般都采用同义词释义的方式。如知网使用的是同义义原释义的方式，它把4000个汉字的所有义项统统列出来，用手工的方法一一观察、挑选、合并同类项，最后提取了2000个义原，如从“医”、“治”、“疗”，提取出“医治”这样一个义原。《现代汉语语义词典》同样采用同义词释义的方法，如对“找”的释义为“寻找”。这种做法的缺陷体现在，一方面，“找”与“寻找”、“医”、“治”、“疗”与“医治”意义上不能完全对等；另一方面，无法揭示词汇复杂的概念结构，这样的义原提取更类似传统词典学中对“基础词汇”的提取，而不是真正意义上的元语言。词典中对词汇单位的释义不是简明就好（如《现代汉语语义词典》的释义方式），释义应该真实反映每个词汇单位所描写的概念结构，如能将概念结构中的构成元素一一指明就更好，这样的释义为动词配价结构的确定提供指定的信息，语义配价的确定也可在释义中找到相应的依据。李葆嘉指出，汉语词典学、词汇学、语义学以及中文信息处理等领域，多年来之所以难以出现重大突破，主要原因在于“尚未对现代汉语元语言系统这一基础性工作进行深入研究”。（李葆嘉 2002：143）无论是面向人的学习词典还是面向机器的语义词典，采用元语言释义是当今词典学释义的发展趋势。

## 2.2 多层次配价结构描写体系

配价结构（диатеза）指特定谓词的语义配价和句法配价之间的对应关系（（1997—2000， т. II: 163），在《组合词典》中莫斯科语义学派使用术语支配模式（модель управления）代替，以表格形式体现其语义配价、深层句法配价和表层句法配价这种多层次配价体系，力求对影响句法、语义对应关系中的各个层级进行细致、准确的描写。语义配价、深层句法配价、表层句法配价三个层级既相互独立，又彼此联系、相互对应，每个层级都有各自的描写手段和规则及主要解决的问题，同时整个配价系统又作为一个整体发挥作用。通常情况下，有多少个语义配价就有多少个深层句法配价，但在实际中深层句法配价的数目可能多于、少于语义配价的数目，即发生配价分裂、配价合并等特殊的句法、语义对应现象。在《组合词典》中，表格第一横栏为深层句法配价（1, 2, 3 …）和对应语义配价变项（X, Y, Z…）。第二横栏是表层句法配价，是深层句法配价在表层句法的体现方式。深层句法配价 1, 2, 3 等是对深层句法结构图示中节点 1, 2, 3 等的抽象。语义配价变项从词汇单位的释义中得出。

下面是俄语动词 учиться（学习）的支配模式（Мельчук 1984： 917），如表格所示：

УЧИТЬСЯ1. X учится Y-у (y Z-a(в W-e))= ‘X 有意识地付出努力获取 (Z 在 W 传授的) Y。’

1=X 谁获取知识或技能	2=Y 力求做到（学会）什么	3=Z 谁（什么）传授知识或技能	4=W 在什么机构
动物名词短语 <sup>⑥</sup>	1. 三格名词短语 2. 未完成体动词不定式 3. на+四格名词短语 4. по+三格名词短语	1. y+二格名词短语 2. под руководством +二格名词短语 3. по+三格名词短语	1. в+六格名词短语 2. на+六格名词短语

汉语配价语义词典中对句法、语义信息大多未分层级描写，一般只注重语义层面的配价框架的描写，而对配价成分在表层中的具体表现形式不予关注，对句法语义对应中的配价分裂、合并等特殊现象更是从不描写。如知网的必要角色框架中只描写词汇语义层面的信息，且语义信息描写的颗粒度较大，例如，buy 的必要角色框架为：<agent, possession, source, cost, beneficiary>，由这个配价框架，我们只知道必须有 agent，这个 agent 表层有哪些表现手段，知网没有给出。《现代汉语语义词典》同样缺少对配价成分在表层中具体表现形式的描写，例如在对“找”进行描写时只指明其配价数、语义角色、语义搭配限制。中国人民大学语言文字研究所编著的《动词大词典》中只给出每个动词的必要配价和语义角色，《中文概念词典》只对动词句子框架作简单描述，例如，动词“吃”的句子框架为“施事+动词+受事”，但显然仅有这些信息是不够的。

词典的最重要的目的是通过对词汇语义、句法信息的标注，最大限度地预测其在句子中的使用情况，这一目的与目前句法—语义界面理论研究目标相一致。《组合词典》中词汇单位的释义揭示词汇单位的语义结构；支配模式表格第一横栏表示语义配价变项 X, Y, Z 和深层句法配价 1, 2, 3 之间的对应关系，它可看成是实现词汇语义和句法接口的工具，是句法、语义接口概念的一种变体和具体应用；第二横栏是深层句法配价在文本中的表达方法。从释义到支配模式第一横栏，再到第二横栏，词汇语义逐步实现了同句法配价的链接。《组合词典》配价结构的这种设置符合学者们对句法、语义界面研究的普遍认识，即从词汇语义要经过词汇语义结构、配价结构、句法结构这几个层级。“通常认为句法—语义界面研究需要考虑三个表征层面，一个是词汇语义结构，一个是论元结构（配价结构），最后一个为句法结构。”（沈园 2007：25）Падучева 在总结“意思 ↔ 文本”模式理论的历史影响时指出，“支配模式是《意思 ↔ 文本》模式理论中最重要的几个概念之一，没有支配模式（配位方式），配价和题元等一些概念，就没有当代的语义学；更为重要的是，在这些概念的基础上产生了解决语义和句法界面问题(проблема соединения семантики с синтаксисом)的机制，即西方语言学中，兴起于上世纪八十年代，至今仍吸引大批学者研究的所谓‘链接问题’(linking problem)。”（Падучева 2004：343）句法—语义界面理论的运用使词典编写更具有理论价值和应用前景。

### 2.3 词汇函数理论

词汇函数<sup>1</sup>是《组合词典》与其他词典相比最具新意的地方。我们知道一个词汇单位与其他单位会有各种组合、聚合关系。各种聚合、组合关系，《组合词典》中用词汇函数表示。词汇函数指一个词汇单位（或其集合）与另一个词汇单位 Y 之间的选择对应关系 f，表示为 Y=f(X)。f 表示特定的抽象概念，如‘极端特征’、‘同义’、‘开始’、‘使不存在’等，用拉丁语缩略形式标记：Magn, Syn, Incep, Liqu 等，如无微不至=Magn（关怀），妻子=Syn（老婆），入睡=Incep（睡觉），取消=Liqu（限制）。X 是词汇函数 f 的自变项，Y 是因变项，因具体 X 的变化而异。例如：《组合词典》中 учиться<sub>1</sub> 的词汇函数有：Syn(учиться<sub>1</sub>)=изучать<sub>1</sub>, учить<sub>2</sub>, брать уроки 等等；S<sub>0</sub>=учеба；Magn=настойчиво, упорно, прилежно, систематически；Incep=идти [учиться]等 30 种。《组合词典》通过词汇函数全面、系统地展现了词汇单位与其他词之间的各种依存关系。词汇函数在词典中的应用可以说是词典学上的一个创举（蒋本蓉 2008：109）

## 2.4 对句子“核心成分”和“非核心成分”进行全面研究与标注

《组合词典》对动词语义结构的描写突破了“配价成分”的限制，实现了对“处所”、“工具”、“时间”等非配价成分的描写。如对 *учиться* 词库中“配价结构”既描写了所谓的配价成分 X（施事），Y（对象），Z（客体），又描写了非配价成分 W（地点），实现对句子“核心成分”和“非核心成分”的全面研究与标注。“因为我们所需要的信息没有核心与非核心之分，信息处理系统对核心和非核心框架元素的操作过程是完全相同的。”（陶明忠，马玉蕾 2007：123）为了比较全面地反映动词对其从属成分的支配能力，为了最大程度上从动词的配价结构去预测其所在句子的结构和意义，配价结构应该既包括施事、受事等核心成分，也包括具有一定的限制性、对语义分析和多义性排除有一定作用的非配价成分，如工具、处所等语义角色。汉语中，有些动词的非配价成分的句法位置灵活，动词的非配价成分也可出现在表层句子的主语、宾语位置，句子中配价成分和非配价成分的句法位置不像俄语中那么固定，更应加强对非配价成分的描写，如动词“吃”，表层中充当宾语的可以是配价成分“饭”，即吃饭；也可以是表示处所、工具的非配价成分“食堂”，“大碗”，即吃食堂，吃大碗；充当动词“跑”的宾语的可以是表示结果的非配价成分，“一身汗”，即跑出一身汗，也可以是表距离、比赛项目的非配价成分，如“跑百米”，跑一百米等，这些非配价成分也应是描写的对象。对非配价成分的描写仅限于那些从动词的词汇意义上可以推导出来，并且是在一定的句法结构的主语、宾语位置中实现了的从属成分。这样，那些几乎可以跟所有的动词共现的时间、处所等成分就应该排除在动词的配价结构之外。

## 3 《组合词典》需要完善之处

《组合词典》编写始于 20 世纪 60 年代，几十年间各种语言学理论层出不穷、语言学知识不断更新，即使是其所涉及的“意思↔文本”模式理论也在不断完善，将新的语言学知识融入到成熟的语言学理论中，可能会使已经比较成熟的语言学理论焕发新的学术生命。

### 3.1 标注语义角色<sup>2</sup>

《组合词典》在语义配价与深层句法配价的对应上，不标注出相应的语义配价的角色类型，这虽然省去了划分语义角色类别的麻烦，但却无助于揭示词汇深层的语义关系。语义角色研究虽面临许多困难，但并没有影响到学者们对它的应用。即使对词汇语义采用谓词分解方法的学者，也不得不将语义角色作为一个派生概念重新纳入到自己的体系中。语义角色在大多数大型词库中多有应用，语义角色标注已成为了语言信息处理中的一个重要课题。目前存在的自然语言词库中，绝大多数仍然选择使用语义角色。加州大学的框架网络项目通过基于场景的语义框架来定义语义角色，框架中的框架元素便是框架网中的角色；知网对于语义角色的策略是尽量挖掘，共采用了 90 个语义角色。（董振东等 2007：7）《中文概念词典》中使用 26 个角色定义动词概念 C 的句子框架（sentence frame）；（于江生，俞士汶 2007：72）关于语义角色对语言信息处理的作用，袁毓林曾举了一个例子说明：鉴于各种语义格的定义和区别的纠缠不清，清华大学计算机系孙茂松教授只对语料标注“因”和“果”两种语义角色；开发了一个 400 万词规模的汉语语义骨架标注语料库，对每个句子标注有面向语义的因事（S）、中心谓词（V）和果事（O）块信息。结果，对于“咬死了猎人的狗”一类歧义格式的消歧起了很好的作用。如此看来，语义角色方面的信息对于语言信息处理是极为重要的，哪怕只是比较粗放的宏观层次上的语义角色。（袁毓林 2007：11）莫斯科语义学派后来的发展中逐渐将语义角色纳入配价结构理论中，如 Апресян 1974 年《词汇语义学》、2004 年《新编俄语同义词词典》和 2010 年《俄语积极词典概况》三本著作中分别归纳使用了 25 种、16 种和 54 种语义角色，Падучева（2004）甚至列出了近百种语义角色，这些都可以为我们进行语义角色标注提供参考。

### 3.2 描写词汇单位的各种语义搭配限制

动词的语义会对题元名词的语义类别有一定的限制，反过来，名词的语义类别，如物质、

行为、非离散性质、活物、自然力、事件等对动词的语义描写有重要作用。《组合词典》对词汇单位的各种语义搭配限制描写得比较笼统、模糊，还没有建立起完整的语义分类系统。名词的分类类别无法开出一张完整的清单，它的数量会随着动词语义规模的扩大而增长。正确的做法是“以述语动词对有关语义角色所施加的语义制约为标准，用动态的描写方法，来建立汉语名词性概念分类体系。”（林杏光 1999：204）汉语界已有多部词典对词汇单位的语义搭配限制进行系统的描写，例如《中文概念词典》“吃”的施事和受事的语义限制分别是“动物；人，人类”和“食物，可吃物”；《现代汉语语义词典》中“找”的2个配价“主体”和“客体”的语义选择限制分别是“人”，和“具体事物”。目前，汉语语义词典中《现代汉语语义词典》和《现代汉语述语动词机器词典》已经尝试建立完整的语义分类体系，《现代汉语语义词典》的语义分类系统可以作为我们描写语义搭配限制的基础。这一系统的有效性和可行性已在《现代汉语语义词典》中得到证实。

### 3.3 增加对“槽关系”的描写

《组合词典》对配价结构的描写还局限在对动词与名词性成分之间的配价关系的描写上，而对名词与名词性成分之间的“槽关系”，即研究语义配价内的偏词和正词之间的语义关系并未涉及。与以英语为代表的西方语言“以动词为中心”不同，汉语不仅是以动词为中心，也是以名词为中心的（这是汉语重要特点之一），所以仅研究动词不研究名词难以满足汉语机器理解的要求。目前清华大学和中国人民大学研制的《现代汉语述语动词机器词典》中对槽关系进行了开拓性的描写。缺少对槽关系描写也是其他语义词典的不足之一，俄语语义词典中至今尚未开始对槽关系研究方面的探索，这或许会为俄语语义词典研究注入新的动力。

### 3.4 增强对不以动词意义为基础的构式的描写

以配价结构为基础的语义词典，存在自身的缺陷，即配价结构从根本上是“动词中心论”的产物，这决定了其在处理语义和句法关系方面的不足：一方面，远非所有语言、所有句式都是以动词为中心的；另一方面，这过分强调了词汇语义对句法结构的决定性作用。目前兴起的构式语法可以弥补这方面的不足，动词语义和句式互动关系是今后研究的方向。构式语法和配价等语法所代表的词汇主义分别代表了两种思潮、两种看待同一问题的不同方式和方法，将二者相互结合、取长补短，共同用来描写句子，实现句子句法和语义的最大限度的联接。词库中尽可能准确地预见动词所能出现的句法环境就要做到既要准确地对动词的论元结构进行描写，又要对不以动词意义为基础的构式进行描写。句法构式具有比较固定的格式，比较容易进行形式化的描写和处理，这给汉语信息处理的进一步发展带来了更多的希望。因此我们认为，把基于词汇的配价结构描写和构式语法结合起来，才能对语义结构和句法结构之间的映射关系进行全面系统的描写，才能更好地为汉语信息处理服务。

## 4 结束语

清华大学和中国人民大学语言文字研究所共建的《现代汉语述语动词机器词典》被专家评价是语言工程上的一个新突破，其研究方法、实现技术、描写内容达到国内外领先水平，论元数目、配价模式、语义角色、语类、句法功能、语义分类、语义特征、论元标记及实例等组成的网格方式对动词进行详细的主要描述。完善后的俄《组合词典》中标注的信息与《现代汉语述语动词机器词典》基本相同，但在元语言释义、配价结构层级性、词汇函数等方面又有自己独特的语言学理论做支撑，如果能在此基础上研发俄汉、汉俄双语语义词典，相信一定会取得不错的效果。

## 附注

1 本文的词汇函数指标词汇函数，标准词汇函数应同时符合两个条件：(a) 该词汇函数建立在数量较多的自变项词汇单位 X 的基础上，即 f 代表的语义范畴要有较高的抽象程度、普遍程度，足以和许许多多的意义单位搭配；(b) 相应的，该词汇函数由数量较多的因变项语言表达式 Y，它们在数量上大体足以分配给组成自变项词汇单位 X 的众多成员。用拉丁语缩略形式标记的均属于标准词汇函数。非标准词汇函数因为不具有概括性，没有进行描写，例如 домашний=囚禁于家中的( арест )。详细内容参见张家骅( 2011 : 195—207 )

2 莫斯科语义学派早期研究中配价既指谓词和名词性成分之间的组配能力，又指语义角色本身。本文将区分这两个概念，用语义角色指代深层语义关系类型。

## 参考文献

- [1] Апресян Ю. Д. Лексическая семантика. Синонимические средства языка[M]. М.: Наука, 1974.
- [2] Апресян Ю. Д. Новый объяснительный словарь синонимов русского языка[Z]. Москва-Вена: Языки славянской культуры, 2004.
- [3] Апресян Ю. Д., Богуславский И.М. и др. Теоретические проблемы русского синтаксиса. Взаимодействие грамматики и словаря[M]. М.: Языки славянских культур, 2010.
- [4] Васильев Л.М. Современная лингвистическая семантика[M]. М.: Высшая школа, 1990.
- [5] Мельчук И.А., Жолковский А.К. Толково-комбинаторный словарь современного русского языка[Z]. Wien: Wiener Slavistischer Almanach, 1984.
- [6] Мельчук И.А. Курс общей морфологии[M]. Т.1-3. М.: ЯРК, 1997-2000.
- [7] Новиков Л.А. Семантика русского языка[M]. М.: Высшая школа, 1982.
- [8] Падучева Е. В. Динамические модели в семантике лексики[M]. М.: Языки славянской культуры, 2004.
- [9] Смирницкий А.И. Объективность существования языка[M]. М.: Издательство МГУ, 1954.
- [10] 董振东, 董强, 郝长伶. 知网的理论发现[J]. 中文信息学报, 2007(4).
- [11] 范晓. 动词的配价与句子的生成[J]. 汉语学习, 1996(1).
- [12] 蒋本蓉. 《现代俄语详解组词典》述评[J]. 辞书研究, 2008(3).
- [13] 李葆嘉. 汉语元语言系统研究的理论建构及应用价值[J]. 南京师大学报(社会科学版), 2002(4).
- [14] 林杏光等. 现代汉语动词大词典[Z]. 北京: 北京语言学院出版社, 1994.
- [15] 林杏光. 词汇语义和计算语言学[M]. 北京: 语文出版社, 1999.
- [16] 鲁川等. 动词大词典[Z]. 北京: 中国物资出版社, 1994.
- [17] 沈园. 句法—语义界面研究[M]. 上海教育出版社, 2007.
- [18] 苏新春, 王惠等. 语言知识库的内化与细化——完善词义知识库之设想[A]. 第六届汉语词汇语义学研讨会[C], 2005.
- [19] 陶明忠, 马玉蕾. 框架网络与汉语信息处理[J]. 语言文字应用, 2007(4).
- [20] 于江生, 俞士汶. 中文概念词典[J]. 中文信息学报, 2007(4).
- [21] 袁毓林. 语义角色的精细等级及其在信息处理中的应用[J]. 中文信息学报, 2007(4).
- [22] 张家骅. 俄罗斯语义学——理论与研究[M]. 北京: 中国社会科学出版社, 2011.

## 《Modern Russian Detailed Combination Dictionary》 and its value for chinese semantic dictionaries

Li Xia

(Centre for Russian Language, Literature and Culture Studies of Heilongjiang University, Harbin 150080,

China)

**Abstract:** 《Modern Russian Detailed Combination Dictionary》 is a thesaurus of the theory "Meaning ⇔ Text" Model of Moscow Semantic School. It is based on the theories of diathesis (government pattern)、lexical functions and metalanguage. The paper try to compare 《Modern Russian Detailed Combination Dictionary》 with chinese semantic dictionaries, analyses its value on the lexical units classification, dictionary definition and diathesis settling, and at the same time, we can see its shortcomings.

**Key words:** semantic dictionary; diathesis; lexical unit; semantic role

**基金项目：**本文系国家社科基金项目基于词汇函数理论的俄汉动词语义世界图景对比研究 ( 13CYY094 )，黑龙江省高校人文社科项目“配价结构与语义词典”( 1252b016 ) 的阶段性成果。

**作者简介：**李侠 ( 1981— )，黑龙江省讷河人。博士，黑龙江大学俄罗斯语言文学语言学与文化研究中心助理研究员。主要研究方向：语义学。

**收稿日期：**2015-06-25

**[责任编辑：张春新]**