

借鉴 TKC 的编写理论建构汉语语义词典的描写框架

蒋本蓉

(黑龙江大学, 哈尔滨 150080)

提 要: 语义词典的建设已成为自然语言处理的基础性工作。世界各国已开发了许多大型的语义词典、词库。本文分析了当今汉语语义词典的主要特点及不足, 系统阐述了莫斯科语义学派“意思↔文本”模式理论的词库——《现代俄语详解组合词典》的编写理论, 借鉴其理论和方法并结合汉语自身的特点, 对其内容进行了补充, 提出了编写汉语语义词典的框架并编写了词条样例。

关键词: 语义词典; 语义元语言; 支配模式; 词汇函数

中图分类号: H14

文献标识码: A

1 引言

语义词典的建设已成为自然语言处理的基础性工作。世界各国已开发了许多大型的语义词典、词库。汉语语义词典的研发经历了不断发展变化的过程, 从以人为对象、以动词词典为主到以机器为对象、系统描写名词、动词、形容词等实词语义信息的语义词典。汉语语义词典的内容与语言学理论结合得越来越紧密, 描写模式也日趋合理完善, 在语义分析、机器翻译、信息提取等领域发挥显著作用。但是到目前为止, 没有一部语义词典能够综合地反映词的语法、语义和句法属性, 这是所有汉语语义词典的一个弱点。

早期的语义词典通常是以动词为描写对象的动词词典。从孟琮等编写的《动词用法词典》到基于格语法理论林杏光等编写的《现代汉语动词大词典》和鲁川的《动词大词典》, 从本质上看, 这些词典是基于动名搭配格式“偏重价量分析”的词典。这种描写方法不能揭示汉语的特点, 因为“汉语中即使是同价动词, 即使一一列举出可能的各种动名搭配格式, 仍然未必能反映汉语大量的特殊句式的动词类别, 如‘把字句’、‘存在句’等”(徐烈炯等 1998: 15)。

董振东创建的“知网”(HowNet)的理论基础是义元分析和语义角色、语义关系的细致描述。知网的核心理论是“知识和意义寓于关系之中, 关系是知识的核心, 关系是知网的灵魂”(董振东等 2007: 4)。知网是充分揭示词汇语义关系的知识词典, 不是全面描写词的语义、句法属性或者说词的聚合关系和组合关系的词汇语义词典。知网重在揭示“知识”(即意义)、类别和“关系”, 词的句法信息显得薄弱。虽然知网借助 90 个语义角色对表示事件类词的必要角色框架进行了细致描写, 但是“知网没有涉及动词的角色变元和句法题元之间的对应关系以及句法题元在表层语句中的分布特征”(薛恩奎 2007: 327)。知网只是静态地标注了动词的语义角色, 没有描写动词的角色变元在句子中的具体表达方法以及使用的语义约束条件。《中文概念词典》(Chinese Concept Dictionary)也存在类似的弱点。虽然它采用 26 个汉语动词论元角色描述了动词的句子框架和论元角色的语义限制, 但它也没有阐述动

词的论元角色在句子中的具体体现。

北京大学的《现代汉语语义词典》(Semantic Knowledge-base of Contemporary Chinese, 以下简称 SKCC) 基于配价理论和格语法相融合的“广义配价模式”(詹卫东 2000: 38) 以词的义项为单位系统描写其语义分类、配价数量、配价成分的语义角色和搭配的语义约束条件。SKCC 的不足与上述词典类似, 没有描写被释词项的句子框架(句型)及在该句型中被释词项使用的语义约束条件, 即 SKCC 只是静态地描写被释词项的语义属性和语义限制, 没有给出在动态的词与词的组合(即句子)中被释词项所具有的属性及搭配限制。此外, SKCC 的语义信息描写得也不够全面。例如, 动词“找”是同形词, “找 A”表示“寻找”, 对“主体”配价的要求是“人”, 对“客体”配价的要求是“具体事物”; “找 B”表示“退还”, 对“主体”的要求是“人”, 对客体的要求是“钱”, 对与事的要求是“人”(王惠等 2006: 140)。实际上“找”的许多组合, 如“你找谁”, “找工作”, “找对象”(男、女朋友)的客体配价既不是“具体事物”, 也不是“钱”。

语义词典应以词的义项为单位, 综合描写语法、语义和句法信息。《现代俄语详解组合词典》(Толково-комбинаторный словарь современного русского языка, 简称 ТКС) 是一部以标题词义项(词项)为对象对语法、语义和句法属性进行集成描写的综合词典。本文将汉语为例, 阐述 ТКС 的编写理论, 借鉴 ТКС 的理论和方法并结合汉语自身的特点, 提出编写汉语语义词典的框架并编写词条样例。

2 ТКС 的理论基础

ТКС 是莫斯科语义学派“意思 \leftrightarrow 文本”模式理论(теория лингвистических моделей «Смысл \leftrightarrow Текст», 简称 ТСТ)的词库。ТСТ 的核心思想是“自然语言被想象为一种转换器, 这种转换器能对应任何对象语的意思列出用自然语言表达的(同义)文本集合, 并能对应任何对象语的文本列出它所表达的(同义)意思集合”(Мельчук 等 1984: 69)。ТСТ 模拟人的语言能力的两个重要方面——从“意思”到“文本”的编码能力和从“文本”到“意思”的解码能力。编码过程就是把语义元语言翻译成自然语言的过程, 而解码则是将自然语言翻译成语义元语言的过程。“意思 \leftrightarrow 文本”模式的双向转换分为 4 个基本层面: 语义、句法、词法和语音。各层次间的转换都有一套规则系统——“意思 \leftrightarrow 文本”模式的语法系统, 而 ТКС 是服务于这个模式转换操作的词汇信息库。

ТКС 的理论基础主要是语义元语言、俄语配价理论和词汇函数理论。

2.1 ТКС 的释义语言——语义元语言

莫斯科语义学派认为, 元语言是被简化和标准化的对象语支语言, 由相对简单的词、句法结构和词法形式构成, 它们的数量比对象语的词汇总量少若干位数。在元语言中不应该有同义词和同音异义词, 一个语言单位只表示一个意义, 一个意义只用一个语言单位表示。元语言的词汇组成包括语义单子(семантический примитив)¹和过渡语义单位。

Апресян (1994: 38) 指出: “如果不能通过词汇语义单位 L_1, L_2, \dots, L_n 来解释词汇语义单位 L , 那么词汇语义单位 L 是语义单子。”即语义单子只参与解读其它语言单位, 自身不能被进一步分解为两个或两个以上其它的语义元素。如 *делать, знать; объект, время; хороший, один; и, или* 等。语义单子是语义元语言的基础。

过渡语义单位的语义比语义单子复杂, 通过有限几个步骤的语义分解能够简化为语义单子。如 *должен, пытаться* 等, P прекратился(停止 P) = начался не P (开始不 P); P продолжается(继续 P) = P не прекращается(不停止 P) ю 在语义元语言中用英文字母 X, Y, Z, P 等表示词项的配价变项。

TKC 遵循以简释繁，避免循环诠释、系统性、全面性的原则，通过对象语语义元语言释文一方面反映某词项的语义、句法属性，另一方面展示该词项与其他词项的聚合关系。TKC 的释义是由词项本身和表示词项的语义配价变项 (X, Y, Z) 组成的表达式。例如，“учиться₁”的释义为“X 在某机构 W 自觉地付出努力，力求获得 Z 传授的知识或技能 Y”。(Мельчук 等 1984: 917) 变项 X, Y, Z, W 是“учиться₁”的语义配价，分别表示主体、直接客体、间接客体(向[按照]谁[什么]学习)、地点。释义解释的不是独立的动词“учиться₁”，而是反映“учиться₁”组合关系的整个述谓表达式，即“X 在 W 向 Z 学习 Y”。这样，释义同时反映语言单位的语义、句法属性。

语义元语言释文为 TKC 中不同词项的相同义素保持相同释文、同类词项保持同类释文提供了保障。例如，对“学习₁”的 5 个近义词 изучать₁ (研究), учить₂ (学习, 背诵), проходить₃ (学习), зубрить (死记硬背), обучаться (学习, 受教育) 的释义分别为“付出努力，获得理论知识，并力求深入”、“付出努力，获得知识，并力求记住”、“付出努力，按照教学机构教学大纲的某具体部分获得知识”、“不加理解地机械学习₂”、“付出努力，系统地获得技能”。释义中有相同的义素“付出”、“努力”、“获得”、“知识”(Мельчук 1999: 129)。语义元语言释文既能保证释义描写的系统性，又能使读者对近义词、反义词等的异同一目了然。

2.2 TKC 语义-句法结构的接口—支配模式 (модель управления) ——俄语配价理论的应用

莫斯科语义学派认为，语义配价是谓词语义单位(动词、形容词、部分名词等)所描写情景的必需参与者，与该谓词释义中的变项相对应。Апресян (1974: 120-21) 指出：“语义配价是词在句法上连接的从属词并与词的释义中每个变项一一对应，语义配价直接来源于词的词汇意义，确定词的语义配价，即确定词所表示情景参与者的数量”。“语义配价”与“句法题元”(семантический акт) 相对应。“句法题元是由文本中出现的与关键词在组合上具有强依存关系的主语和补语决定的”(Мельчук 1974: 135)。也就是说，语义配价是深层词汇语义结构中的变元，与元语言释文中的变项 (X, Y, Z 等) 相对应；句法题元则是表层句法结构中的常项，由文本中实际出现的与关键词具有强依存关系的主语和补语决定。

TKC 采用支配模式描写被释词项的语义配价和句法题元间的关系以及限制句法题元表达方式的语法、语义条件，支配模式是描写谓词语义与句法结构的接口。支配模式体现词项的语义配价框架和每个语义配价所有可能存在的表达方法，以表格形式呈现。表格的每一个纵列是一个语义配价，而每一个横行是相应的句法题元在表层体现的一种方法。例如，“учится₁” (动词) 的支配模式如表 1 所示：

1 = X [主体]	2 = Y [客体 1]	3 = Z [客体 2]	4 = W [处所]
1. 名词 (名词短语)、代词 1 格	1. 名词 3 格 2. 动词 (动词短语) 3. на+名词 4 格 4. по+名词 3 格	1. у + 名词 2 格 2. под руководством +名词 2 格 3. по + 名词 3 格	где

“学习₁”有 X、Y、Z、W 4 个语义配价。X 在表层句法结构中体现为名词 (名词短语) 或代词。Y 在表层句法结构中有 4 种表达方法：①表示“学习什么”：учиться музыке；②表示“学习做什么”：учиться плавать；③表示“学习成为从事某职业的人”：учиться на врача；④表示“学习某课程”：учиться по физике。Z 在表层结构中有 3 种表达方式：①表示“向 (跟) 某人学习”：учиться у учителя；②表示“在某人的指导下学习”：учиться под руководством профессора；③表示“按照什么学习”：учиться по книге。W 在表层中体现为表示地点的介词词组。TKC 支配模式同时体现了被释词项的语义配价和句法题元。

在支配模式的下方是“支配模式的限制”。支配模式的限制，即句法题元表达方式的限制，指的是被释词项的句法题元在语法、语义等方面兼容共现的限制条件。确定某题元 D_{ij} 的方法为以 i 代表支配模式表中的纵列号码， j 代表横行号码。例如，“учиться₁”的支配模式限制是：

- ① 1.1 为动物名词。
- ② 2.3 为实践性、操作性强的职业。如 учиться на врача (на шофера)。
- ③ 2.4 的主体通常为学生，在 Y 中有评价性的补语。如 По физике он учится неплохо。
- ④ 3.1, 3.2 只用于人。
- ⑤ 3.3 表示信息载体(书、电脑光盘等)，3.3 可与 3.1 或 3.2 同现。
- ⑥ 如果 W 表示某个学校，并且 Y 和 Z 不出现， W 可以没有具体的定语，Сын учится 表示“儿子是某学校的学生”。

2.3 ТКС 描写词汇组合的词汇函数理论

词汇函数是 ТКС 与其他词典相比最具新意之处。“意思 \leftrightarrow 文本”模式中用词汇函数来描写一个词项与其相关词项之间的依存关系。Мельчук (1999: 101) 指出， f 描写一个词项 X 与另一个词项或词项的集合之间的依存关系 $\{Y_i\} = f(x)$ 。 f 是词汇函数的名称项，表示特定的抽象语义类型，如“同义”、“反义”、“极大特征”、“开始”等，用相应拉丁语的缩略形式 Syn, Anti, Magn, Incep 等表示。 X (关键词或词组) 是词汇函数 f 的自变项， $\{Y_i\}$ 是词汇函数 f 的因变项。

被释词与其相关词的关系有两种——自由搭配关系和固定搭配关系。具有自由组合关系的词汇函数是常值词汇函数，从描写词汇搭配的角度看，没有描写的必要。词汇函数通常指变值函数，即反映熟语性的固定搭配关系。例如，词汇函数 Magn 表示被释词的“极大特征”(非常)：крепкий = Magn (чай)，同样，проливной/дождь, густой/волосы, гробовой / молчание, крепко/спать, полностью / соглашаться 等。再如，词汇函数 Oper 表示某种概括性的动作“实行”：делать = Oper₁ (вывод)，оказывать / помощь, принимать/решение, вести / борьба, заниматься/исследование, совершать / ошибка, быть в / растерянность。

自然语言中熟语性搭配的数量数不胜数，但它们可以概括为几十个数量有限的函数类型，如 Syn (同义)，Anti (反义)，Conv (换位义)，Incep (开始)，Cont (继续)，Fin (结束)，Caus (致使)，Liqu (使不存在)，Ger (构词派生)，Gener (上位概念)，Son (典型声响) 等。目前所研究的词汇函数有 70 余种。词汇函数理论以形式化方法系统、深入地揭示了词与词之间的抽象语义关系类型。

3 建构汉语语义词典的描写框架及词条试编

我们借鉴 ТКС 的理论和方法，并根据汉语自身的特点对 ТКС 的内容进行了补充，提出了建构汉语语义词典的框架。

3.1 汉语语义词典的描写框架

汉语语义词典(以下简称词典)总的描写原则是——以词的义项为单位，对其语法、语义、句法等属性进行综合描写。具体原则如下：

1) 参照 SKCC 对被释词的拼音进行标注。声调用“1, 2, 3, 4, 5”表示，其中“5”表示轻声。例如，“希望”的拼音是“xi1wang4”，“儿子”的拼音是“er2zi5”。

2) 用英文缩写字母标注被释词的词性。名词、动词、形容词等分别用相应的英文缩写形式表示为“n, v, a”等。

3) 给出被释词的修辞特征。除了被释词本身，还要描写被释词不同义项上的同义词、反义词的修辞色彩。用(中)、(书)、(口)、(方)等分别表示中性词、书面语、口语和方言等。

4) 参照 SKCC 的分类标准, 描写词的语义类别。这是对 TKC 内容的补充。被释词不同义项的语义类别可能会不同。例如, “学校 1” 表示“教育机构”, 语义类别是“机构”; “学校 2” 表示“校舍”, 语义类别是“建筑物”; “学校 3” 表示“全体师生”, 语义类别是“团体”。

5) 词典释义的对象不仅仅是词项本身, 而是由词项和表示词项语义配价的变项 (X, Y, Z) 组成的表达式。例如, 词典释义的不是“怕”, 而是“X 怕 Y”。词典采用 TKC 的释义原则对词义进行注释。

6) 必要时在“注释”中给出与关键词有关的联想意义、同义词的意义辨析或百科知识。

7) 词典中词项用阿拉伯数字“1, 2, 3”等表示, 如“治疗 1”, “治疗 2”等; 用英文字母“A, B, C”表示同形异义词, 如“解决问题 A”(表示“完成事情、任务”)和“解决问题 B”(表示“起作用, 有效”, 如“这种药治疗咳嗽很解决问题”)。

8) 在支配模式中描写被释词的语义配价和相应句法题元的表达方法。语义配价也被称为语义角色、格关系, 由于语义角色的数量及定义不统一, 我们在词典中采取一种简化的处理办法。我们把语义配价概括地分为“主体”、“客体”、“方式”、“处所”、“缘由”等。

9) 在支配模式中描写被释词句法题元所作的句子成分。句子成分的描写是对 TKC 内容的补充。俄语是屈折语, 主要用词尾变化表示句法关系, 因此 TKC 在支配模式的句法题元中标明名词短语的格形式。汉语是分析语, 主要用词序、虚词表达句法关系, 所以在汉语语义词典支配模式的句法题元中要标明名词短语的句法成分和必需的语法助词。汉语句子成分有固定的词序位置, 标句子成分就等于标词序。

10) 利用 TKC 的词汇函数理论描写汉语的熟语性固定搭配。除了词与词之间的固定搭配, 词典还描写词的内部词素与词素的熟语性组合。例如, “热爱”是一个独立的词, 不是词组, 但“热”是表达“爱”的熟语性组合词素, 表示“爱的强烈程度”(Magn)。实际上, 汉语中有许多由熟语性制约关系的词素构成的词, 如“执教”、“从政”、“行医”、“打工”、“经商”等词的划线部分都表示一种概括性的“实行”(Oper), 我们把组合性熟语词也纳入到词典中, 必要时在句法题元表达方式的限制栏目中标出与被释词素组合的是词素, 而不是词。

11) 熟语的描写原则与词的描写原则相同, 也以义项为单位。例如, “酒香不怕巷子深”列入“怕”的第一个义项中, 表示“害怕”; 而“人怕出名, 猪怕壮”则列入“怕”的第二个义项中, 表示“对……有害”。

12) 词典给出被释词使用的具体例证。例证分为两部分: 第一部分是句法题元表达方式的例证, 即支配模式表格给出被释词的句法题元表达方式后, 词典给出具体的例子来加以说明; 第二部分是词汇函数组合搭配的例证, 即在给出词汇函数表达式的意义取值后, 举例说明。

13) 词典对部分概念的界定。限于篇幅, 关于词汇函数的所有类型可参见《俄罗斯当代语义学》(张家骅 2002, 2003), 在此不作赘述。

有些关键词的熟语性搭配属于非标准词汇函数, 没有相应的词汇函数表达符号, 非标准词汇函数的意义用自然语言进行描写并用 F_1, F_2 等表示出来。例如, “没及时治疗”表示为 $F_1 =$ 耽误治疗, 误诊 2。F 的下标表示非标准函数的先后次序。

本文参照生成语法的做法, 支配模式表格中的名词短语包括单个的人称代词、名词; 动词短语包括单个动词。

4.2 词条样例

限于篇幅, 我们只以“治疗”一词为例。为了便于读者理解, 对词条中出现的词汇函数我们在括号内用自然语言加以解释。

治疗

拼音: zhi4liao2

词性: v

修辞色彩: 中性

语义类别: 行为

释义:

治疗 1: X 用 Z 给 Y 治疗 W= X 用方法、手段 Z 作用于 Y, 目的是使 Y 的疾病、外伤 W 不存在。

支配模式:

1 = X [主体]	2 = Y+ W [客体 复合语义配价]		4 = Z [方式]
1.名词短语—主语	1.领属名词短语—宾语		1. “用+名词短语” — 状语
2.名词短语—主语	2. “给+领有名词短语” — 状语 (由上面的复合配价分裂)	3.所属名词短语—宾语(由上面的复合配价“他的病”分裂)	2. “用+名词短语” — 状语
3. “名词短语+治疗” — 主谓谓语	4.领属名词短语—主语		3. “用+名词短语” — 状语

句法题元表达方式的限制:

1) 表格句法题元栏中的三个纵栏表示“治疗 1”的三个句法题元框架。

2) 1.1, 1.2 为表人的动物名词短语。

3) 患者 Y 和身体部位的疾病、外伤 W 是领属关系, 在语义层面只能是一个语义配价, 即复合语义配价 Y (患者)+W (疾病、外伤), 在表层, 即在句法题元层面, 可以分裂为不同的句法题元, 也可以由一个复合题元表示。

4) 2.1=Y+W 为领属名词短语, 二者是整体与部分的关系, 可表示: (1) 患者的疾病; (2) 疾病; (3) 患者。

5) “客体”复合语义配价可分裂为领有者和所属事物两个部分, 这两个部分本文分别用“领事”/“属事”表示。

2.2 为动物名词短语, 是“客体”复合语义配价分裂为“领事”对应的句法题元之一, 可以说, “医生给病人 (或熊猫) 治疗消化不良”, 不能说*“我们给小草治疗疾病”。

6) 2.3 是“客体”复合语义配价分裂为“属事”对应的句法题元之一。

例证:

1) 医生用某种药物治疗他的胃溃疡; 老中医用针灸治疗关节炎; 这个医生专门治疗风湿病; 他只治疗危重病人。

2) 医生用某种药物给病人治疗胃溃疡; 医生给他治疗心脏病。

3) 他的胃溃疡医生在用药物治疗。

词汇函数:

Syn (同义词): (中) 医治、诊治、医疗; (口) 看病 1²、治 (任何疾病都应该赶快医治; 这家诊所免费为孤寡老人诊治; 他没有发生过一起医疗事故; 大夫没在家, 他给人看病 1 去了; 你有病就趁早治)。

Conv (换位词, 客体作句子的主语): (中) 就医、看医生; (书) 就诊、求诊、求医; (口) 看病 2 (生病时应及时就医; 你不舒服, 就别来了, 去看医生吧; 现在很方便, 病人可以到社区卫生院就医; 病人因患急性肠炎到医院求诊; 他到处求医, 无法治愈; 我到医院看病 2 去)。

S₁ (第 1 题元): 医生、大夫。

BonS₁ (Bon 表示事物、行为的良好特征): 神医。

FigurBonS₁ (Figur 表示隐喻): 扁鹊、华佗 (这位医生简直是华佗再世)。

AntiBonS₁ (Anti 表示反义): 庸医、巫医、江湖医生。

S₂ (第 2 题元): 病人、患者 (通常为人)。

S₃ (第 3 题元): 疾病、外伤。

S₄ (第 4 题元): 药物、疗法 (如物理、化学疗法)、手段、工具等。

S_{loc} (处所): 医院、诊所、卫生院、保健院。

AntiAble₃ (Able 表示可能性特征, 3 表示第 3 题元—疾病的词作主语): 不治之症 (艾滋病是不治之症)。

Perf (表示行为、动作的完成): (中) 治好、治愈、根治; (口) 看好 (医生治好了他的胃病; 医生治愈了许多患者; 鼻炎不能根治; 大夫把我的病看好了)。

Ver (表示事物、行为应该具有的特征): 及时 (幸亏医生治疗及时, 病人脱离了危险)。

AntiVer: 误诊 1 (由于医生误诊, 病人死了)。

Oper₁ (表示概括性的动作“实行”, “治疗”的第 1 题元作句子的主语): 进行治疗 (医生对患者进行治疗)。

Oper₂ (表示概括性的动作“实行”, “治疗”的第 2 题元作句子的主语): 接受治疗 (病人需要住院接受治疗)。

F₁ = 没及时治疗: 耽误治疗、误诊 2 (病人没及时就医, 耽误了治疗; 病人离医院很远, 因而误诊 2)。

关于疾病和治疗疾病的科学: 医学、兽医学。

治疗 2: Z 能 (可以) 治疗 W = 方法、手段、工具 Z 能够使疾病或外伤 W 不存在。

支配模式:

1 = Z [方式]	2 = W [客体]
1. 名词短语—主语	1. 名词短语—宾语
2. 动词短语—主语	

句法题元表达方式的限制:

- 1) 1.1 是非动物名词。
- 2) 2.1 的名词短语仅表示疾病、外伤。

例证:

- 1) 板蓝根能治疗感冒; 这种新药能治疗重感冒; 针灸可以治疗耳鸣; 频谱治疗仪能治疗多种疾病。
- 2) 拔火罐能治疗腰背痛; 赤脚在沙地上行走可以治疗风湿。

词汇函数:

Syn: (中) 医治、消除、对……有效; (口) 治。

Mag_n (表示“极大特征”): 彻底 (手术能够彻底治疗这种病)。

S₁: 药物、疗法、手段、工具等。

A₁ (表示“具有……特征的”): 能治病的 (人们发现, 野菜是一种能治病的良药)

治疗 3: Y 在 X 那里用 Z 治疗 (自己的) W = Y 让职业医师 X 用 Z 治疗 Y 的疾病 W。

支配模式:

1=Y+ W [客体复合语义配价]		2 = X [主体]	3= Z [方式]
1. 领属名词短语—主语	2. 所属名词短语—宾语	1. 在+名词短语+那里 (那儿) —状语	1. (采) 用+名词短语—状语

句法题元表达方式的限制:

- 1) 1.1 的领属名词短语是复合客体语义配价分裂为领事对应的句法题元。
- 2) 1.2 的所属名词短语是复合客体语义配价分裂为属事对应的句法题元。

例证:

1) 我在张医生那里用中药治疗乙型肝炎; 许多患者在我们这儿采用手术治疗近视, 效果很好。

2) 朋友在一位著名专家那儿治疗过敏性鼻炎; 他没上课, 在治疗喉炎。

词汇函数:

$S_1, S_3, S_4 =$ 治疗 1 的 S_2, S_1, S_4 。

Syn: (中) 接受治疗 (Oper₂); (口) 治;

(短时体) (口) 治一治 (我想治一治这个老毛病)。

Perf: (中) 治好、治愈; (口) 看好 (我在一位老中医那里治好 (治愈、看好) 了皮肤病)。

5 结语

虽然 TKC 是上世纪 80 年代出版的词典, 但它融合了现代语义学和句法学的研究成果, 其编纂原则和结构设置对今天的语义词典编写依然具有指导意义。本文借鉴 TKC 的理论和方法, 结合汉语自身的特点, 对 TKC 进行了补充, 提出了编写汉语语义词典的框架。我们以词项为单位, 按照 TKC 的释义原则采用对象语语义元语言进行释义, 系统化、标准化地描写熟语, 综合描写词的语义属性和句法属性, 采用形式化的词汇函数语言描写词与词的熟语性组合, 同时补充描写了词的语义类别、修辞色彩和句子成分及必需的语法助词。我们相信, 本文提出的语义词典编写的新模式对汉语知识库的建设会提供新的思路。

附注

1 目前对该术语 (英文 semantic primitive) 有多种译法: “语义基元”, “语义原词”, “语义原语”, “语义原子”, “语义公因数”, “语义单子”, “语义原生意义” 等, 本文采用 “语义单子” 这一译法。

2 “看病 1” 的语义配价结构中已经包含 “客体” —— 疾病, 因此不能再和表示疾病的名词短语连用。例如, 不能说 *大夫给人看病哮喘。

参考文献

- [1] Апресян Ю. Д. 1974 Избранные труды, том I. Лексическая семантика [M]. М., Наука.
- [2] Апресян Ю. Д. 1994 О языке толкований и семантических примитивах [M]. Изв. АН СССР. Сер. Лит. и яз., №4.
- [3] Мельчук И. А., Жолковский А. К. 1984 Толково-комбинаторный словарь современного русского языка [Z]. Вена, Wiener Slawistischer Almanach Sonderband.
- [4] Мельчук И. А. 1974 Опыт теории лингвистической модели «Смысл ↔ Текст» [M]. М., Наука.
- [5] Мельчук И. А. 1999 Опыт теории лингвистической модели «Смысл ↔ Текст» [M]. М., Языки русской культуры.
- [6] 董振东 董强 郝长伶 2007 知网的理论发现 [J], 中文信息学报, 第 4 期。
- [7] 林杏光 王玲玲 孙德金主编 1994 现代汉语动词大词典 [Z], 北京: 语言学院出版社。
- [8] 鲁川 1994 动词大词典 [Z], 北京: 中国物资出版社。
- [9] 孟琮 郑怀德 孟庆海 蔡文兰 1987 动词用法词典 [Z], 上海: 上海辞书出版社。

- [10]王 惠 詹卫东 俞士汶 2006 “现代汉语语义词典”的结构及应用[J], 语言文字应用, 第 1 期。
- [11]徐烈炯 沈阳 1998 题元理论与汉语配价问题[J], 当代语言学, 第 3 期。
- [12]薛恩奎 2007 И.А.Мельчук 的“意思↔文本”学说[J], 当代语言学, 第 4 期。
- [13]詹卫东 2000 基于配价的汉语语义词典[J], 语言文字应用, 第 1 期。
- [14]张家骅 2002 词汇函数的理论和应用[J], 外语学刊, 第 4 期。
- [15]张家骅 2003 俄罗斯当代语义学[J], 北京: 商务印书馆。

A Framework of Chinese Semantic Dictionary Based on the Theories of ECD

JIANG Ben-rong

(Heilongjiang University, Harbin 150080, China)

Abstract: Semantic Dictionary is an essential component in Natural Language Processing. Now many Semantic Dictionaries and Lexical Databases have been developed. This paper analyses the merits and defects of Chinese semantic dictionaries and then systematically expounds the theoretical foundation of *An Explanatory Combinatorial Dictionary of Modern Russian* (ECD), namely the lexicon within the framework of the Meaning-Text Model (MTT) of Moscow Semantic School. Finally a framework of Chinese semantic dictionary is presented on the basis of theories and methods of ECD and some entry-samples are compiled.

Key words: semantic dictionary; metalanguage; government Model; lexical functions

收稿日期: 2009-08-12

基金项目: 本文系教育部人文社会科学重点研究基地重大项目“<意思↔文本>模式的自动转换机制——方法和词库”(批准号: 05JJD740006)的阶段性研究成果。

作者简介: 蒋本蓉(1973-), 女, 黑龙江哈尔滨人, 黑龙江大学俄语学院讲师。主要研究方向: 俄语语义学, 词典学。

[责任编辑: 彭玉海]