

支配模式在汉语语义词典中的应用*

蒋本蓉

(黑龙江大学, 哈尔滨 150080)

提 要:《现代汉语语义词典》(Semantic Knowledge-base of Contemporary Chinese, 简称 SKCC)是一部机器可读的汉语语义词典。本文概括介绍 SKCC的内容、特点并分析其主要缺点,尝试用莫斯科语义学派“意思↔文本”模式中的“支配模式”来克服 SKCC的不足。支配模式同时体现关键词用于所有义项时的语义、句法信息以及词义搭配的语义限制,应成为语义词典知识描述系统不可缺少的组成部分。

关键词:语义词典;支配模式;语义配价;句法配价

中图分类号: H16

文献标识码: A

文章编号: 1000 - 0100(2009)04 - 0129 - 3

Mode of Government in Semantic Dictionaries of Chinese

Jiang Ben-rong

(Heilongjiang University, Harbin 150080, China)

Semantic Knowledge-base of Contemporary Chinese (SKCC) is a Machine Readable Dictionary. This paper has discussed the contents, character, analyzed the shortage of the SKCC and put forward a settlement — the Mode of Government of system “Meaning Text” of the Moscow semantic school. The Mode of Government presents the semantic and syntactic information of all the senses of a word and semantic restraint in word combinations. The Mode of Government should be a necessary part of the Semantic Dictionaries.

Key words: semantic dictionary; mode of government; semantic valence; syntactic valence

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、以动词为主,类似德语、法语基于动名搭配形式偏重价量分析的配价词典,如孟琮等(1987)编写的《动词用法词典》,林杏光等编写的《现代汉语动词大词典》(1994)和鲁川编写的《动词大词典》(1994)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。语义词典的描写模式日趋成熟。

从1994年开始,北京大学与中国科学院计算所着手研制面向汉英机器翻译的《现代汉语语义词典》(Semantic Knowledge-base of Contemporary Chinese, 简称 SKCC),其目的是“在语法分析的基础上为计算机自动分析汉语句子和生成英语句子提供更深入的语义信息”。(王惠等2003: 352)本文以该词典为对象,概括介绍其内容、特点,分析其缺点,并尝试用莫斯科语义学派“意思↔文本”模式中的“支配模式”弥补 SKCC的不足之处,在此基础上

提出建构汉语语义词典的一个更为合理的语义知识描述框架。

1 《现代汉语语义词典》概述

SKCC主要包括两方面的内容——词的语义分类和语义搭配。一方面,SKCC对名词、动词、形容词等实词进行了系统的语义分类;另一方面,SKCC以词的义项为单位全面描写词的语法、语义信息,包括拼音、词类、兼类(名词“锁”兼为动词)、同形(字形、词类相同但词义不同,如“抄近道”的“抄”与“抄作业”的“抄”)、义项、释义、语义类、配价数量、配价成分的语义角色(主体、客体、与事)和搭配的语义约束(即主体、客体、与事所属的语义类名称,例如,“死”的主体语义约束条件是“生物”)等。其中词的语义搭配(后三项)是描写重点。

SKCC的建构原理被称为“广义配价模式”,即基本采

* 本文系国家社会科学基金项目“建构《俄汉平行对照语义关系词典》理论与实践”(09BYY066)的阶段性研究成果。

用配价理论的模式并融合格语法的语义角色系统。广义配价模式包括语义分类、配价、论元角色选择限制、配价成分变化四个方面(詹卫东 2000: 40)。其中对配价成分变化的描写只是 SKCC 初期的尝试,广义配价模式实际上只包括前三个方面。

SKCC 把格语法的论元概念从动词扩展到谓词(即动词、形容词和名词),并把格系统进行简化,“主体”和“客体”的含义更加广泛,“主体”包括通常的“施事”、“当事”,还包括部分名词和形容词的配价成分。动词的论元有 6 个:“主体”、“客体”、“与事”、“工具”、“处所”为核心论元,“空间”、“时间”为外围论元。对大多数动词只标注前 3 个核心论元,其他论元必要时才标注。SKCC 把格语法和配价理论融合在一起,这样,论元成分的数量多于配价的数目。例如,动词“找”是一个同形词,当“找”表示“寻找”的意义时,其语义类别为“对待”,是 2 价动词,即可以搭配“主体”和“客体”配价,对“主体”配价的要求是“人”,对“客体”配价的要求是“具体事物”,如“他找书”;当“找”表示“退还”意义时,其语义类别也是“对待”,但却是 3 价动词——“主体”、“客体”、“与事”,对“主体”的要求是“人”,对客体的要求是“钱”,对与事的要求是“人”,如“我找他 3 块钱”(王惠等 2006: 140)。

SKCC 采用“广义配价模式”系统描写词的语义分类和语义搭配。总的来看,SKCC 对词的语义知识描写是比较全面、深入、合理的。与以往的语义词典相比,SKCC 在理论基础、内容、规模、描写框架等方面更加成熟,更加合理。实践证明,SKCC 对于汉英机器翻译系统的句义分析、词汇歧义消解、提高翻译精度方面有较明显的作用。

2 SKCC 的缺陷及解决办法

SKCC 的语义知识描述框架并非完美无缺。范晓曾指出,理想的动词配价词典是“某种语言生成句子的词典”,应描写述谓结构的性质(动作、性状、变化、关系等)和所联系的语义角色以及某些经常联系的说明语,在此基础上,对该动词的形式特征以及它组合成的述谓结构所生成的句模和各种可能有的句型进行详尽的描写。如果可能,还可标注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)。

范晓的观点恰好说明了 SKCC 最大的缺陷——SKCC 没有以词的义项为单位全面、系统描写句型并且没有详尽刻画该义项在不同句型中使用的语义约束条件,即 SKCC 描写的词的组合信息是静态的,不是动态的。例如,上文中“找”的第 1 义项的语义信息是不全面的,许多组合,如“你找谁”、“我找你们领导”、“找工作”、“找对象”(男、女朋友)等的客体配价并不是“具体事物”,如果不全面标注“找”在所有不同句子中的组合搭配信息及搭配限制,计算机就无法译出正确的句子。

我们认为,采用莫斯科语义学派“意思 \leftrightarrow 文本”模式的词库《现代俄语详解组合词典》(

简称 TKC) 中的“支配模式”可以克服 SKCC 的不足。支配模式表示被释词项的语义配价和句法题元间的对应关系以及限制句法题元表达方法的语法、语义条件,支配模式是描写谓词语义与句法结构的接口。按照莫斯科语义学派的观点,语义配价直接来源于词汇语义单位的词汇意义,与词汇语义单位的元语言释文中表示必须情景参与者的变项一一对应,而句法题元是由文本中实际出现的与关键词具有强支配关系的主语和补语决定的,即谓词在句法结构层面上连接的句法位(

1974: 120 - 121, 1974: 135)。在支配模式中对应关键词的每个语义配价给出相应句法题元的表达方法;反过来,对应关键词的每个句法题元在支配模式中给出语义解释,即指出关键词所示的情景中该题元表示的角色(

3 支配模式

支配模式分为“支配模式表格”和“支配模式限制”两个栏目。

3.1 支配模式表格

支配模式以表格形式呈现。表格中每一个纵列对应被释词项的一个语义配价,而每一个横列描写相应句法题元的一种表达方式。例如,动词“治疗 1”的释义为“X 采用方法 Z 作用于 Y,目的是停止 Y 的疾病、创伤 W 或是停止 W 的后果”(TKC 1984: 400)。《现代汉语词典》(简称《现汉》)2006: 1758 对“治疗”的释义为“用药物、手术等消除疾病”,从释义中可以看出,俄、汉语中“治疗”有 4 个语义配价:主体 X(医生),客体 Y(病人)、W(疾病、外伤)和方式 Z。其支配模式如下图所示:

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	2. “用 + 名词短语”—状语
3. “名词短语 + 治疗”—主谓谓语句	3. 所属名词短语—宾语	3. “用 + 名词短语”—状语
	4. 领属名词短语—主语	

患者 Y 和身体部位的疾病、外伤 W 是领属关系,在语义层面是一个复合语义配价——“客体”,在表层,即句法题元层面,可以由一个复合题元表示,即由领属名词短语作宾语,如“老中医用针灸治疗病人的关节炎”,也可分

裂为不同的句法题元,即分裂为领有者和所属事物两个部分,可表示:(1)患者的疾病;(2)疾病;(3)患者。例如,“老中医用针灸给病人治疗关节炎”;“这个医生专门治疗风湿病”;“他只治疗危重病人”。汉语中有许多类似的配价分裂现象,如“王冕死了父亲”,“他摸我头”,“电视降价了”等。此外,主体 X 还可以用“名词短语+治疗”表示,客体复合语义配价用领属名词短语表示,在句中作主谓谓语,如“他的胃溃疡医生在用药物治疗”。

除了配价分裂,支配模式还能反映配价合并现象。配价合并指的是谓词的两个语义配价用一个句法题元体现。例如,兼语句动词“佩服”可以把客体配价和缘由配价合二为一,如“我佩服你(客体)的勇气(原因)”;“父亲佩服敢想敢干的年轻人(原因+客体)”;“所有人佩服他(客体)门门功课得100分(原因)”。“佩服”的支配模式简单表示为下图,其中第三行跨两个纵列,共同体现客体和缘由:

1 = X(主体)	2 = Y(客体)	3 = (缘由)
表人名词、名词短语、代词	表人名词、名词短语、代词	1 动词或形容词短语 2 小句
	1. 领属复合名词短语 2. 修饰复合名词短语	

3.2 支配模式限制

“支配模式限制”栏目描写关键词的句法配价兼容共现的约束条件。“治疗 1 的支配模式限制有 2 个:(1)主体为表人的名词短语。(2) 2 为动物名词短语,不表示器官。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”。可以说,“医生给他治疗痛风”,“医生给老王治疗手部骨折”,不能说*“医生给他的脚治疗痛风”;*“医生给老王的手治疗骨折”。

“治疗 2 的释义为“Z 治疗 W = 方法或手段 Z 能够停止疾病或创伤 W (Y 的)或 W 的后果”(TKC 1984: 403),当表示方法、手段的词作主语时,“治疗”有三个语义配价:方法 1Z 客体 2W (什么被停止)和主体 3Y (谁生病/受伤),其中主体隐匿在表层句法结构中,只体现在深层句法结构中。Z 和 W 在表层句法结构中都用名词或名词短语表示。例如,“板蓝根治疗感冒”;“针灸治疗颈椎病”。

“治疗 3 表示“Y 让职业医师 X 用 Z 治疗 Y 的疾病 W”;“治疗 3 有 4 个配价:1Y (谁接受治疗), 2W (治疗什么), 3X (谁治疗), 4Z (用什么治疗)。其中 Y 在表层句法结构中有三种表达方法——名词、名词短语和代词;W 用名词或名词短语表示,是必有配价;X 用“在+名词(名词短语、代词)表示,Z 用“(采)用+名词短语”表示。例

如,“我在张医生那里用中药治疗腰间盘突出”;“许多患者在我们这儿采用手术治疗近视,效果很好”。

4 结束语

我们认为,在 SKCC 内容的基础上补充“支配模式”的描写是编写汉语语义词典的一个更合理的语义知识描述框架。在机译系统中,语义词典的核心内容是具体地、详尽地反映词的语法、语义和句法信息。词的语义和句法信息是相互融合的,脱离句子孤立、静态地描写词的语义信息是不行的,对于以机器为对象的语义词典来说,“支配模式”应是不可缺少的部分。支配模式同时描写关键词用于所有义项时的语义与句法信息,包括语义配价、句法配价、语义角色,列出含关键词的句型并结合实例详细刻画词义组合的搭配限制,从而能辅助机器综合地、正确地分析判断,哪些句子是正确的,哪些句子是错误的,有效地进行词义消歧。支配模式对人们学习语言以及计算机处理自然语言都有实用价值,缺少了“支配模式”信息的语义词典是不完整的,不是真正意义上的语义词典。

参考文献

范晓. 动词的配价与句子的生成 [J]. 汉语学习, 1996 (1)

林杏光 王玲玲 孙德金. 现代汉语动词大词典 [Z]. 北京: 北京语言学院出版社, 1994

鲁川. 动词大词典 [Z]. 北京: 中国物资出版社, 1994

孟琮 郑怀德 孟庆海 蔡文兰. 动词用法词典 [Z]. 上海: 上海辞书出版社, 1987

王惠 詹卫东 俞士汶. “现代汉语语义词典”的结构及应用 [J]. 语言文字应用, 2006 (1).

王惠 俞士汶 詹卫东. 现代汉语语义词典 (SKCC) 的新进展 [A]. 语言计算与基于内容的文本处理 [C]. 全国第七届计算语言学联合学术会议, 2003

詹卫东. 基于配价的汉语语义词典 [J]. 语言文字应用, 2000 (1)

中国社会科学院语言研究所词典编辑室. 现代汉语词典 [Z]. 商务印书馆, 2006

[M]. .. : - , 1974

“ ↔ ” [M]. .. : , 1974

[Z].

., Wiener Slawistischer A manach Sonderband, 1984

