

文章编号: 1003-0077(2022)08-0029-08

《动词句法语义信息词典》知识体系及其检索界面

袁毓林¹, 曹宏²

(1. 澳门大学 人文学院中国语言文学系, 澳门;
2. 北京大学 考古文博学院, 北京 100871)

摘要: 该文首先介绍《动词句法语义信息词典》的体系结构与理论背景; 然后, 介绍该词典所区分的 8 种动词小类及其定义; 重点介绍该词典为动词所设置的 22 种语义角色及其定义, 由这些语义角色的不同配置而造成的 20 来种句法格式及其例句, 及其所考察的动词的 9 种主要的语法功能及其对于该词类的隶属度; 最后, 给出该词典的检索系统的界面截图, 交代其相应的纸质版本的情况。

关键词: 动词句法语义信息词典; 动词小类; 语义角色; 句法格式; 语法功能
中图分类号: TP391 文献标识码: A

An Introduction to the Syntactic-Semantic Knowledge-Base of Chinese Verbs

YUAN Yulin¹, CAO Hong²

(1. Department of Chinese Language and Literature, University of Macau, Macau, China;
2. School of Archaeology and Museology, Peking University, Beijing 100871, China)

Abstract: This paper introduces the contents and structure of *The Syntactic-Semantic Knowledge-Base of Chinese Verbs* (KB@verb). First, the structural system and theoretical foundation of KB@verb are introduced. Secondly, KB@verb classifies verbs into eight sub-classes and defines 22 semantic roles, which are configured into decades of syntactic formats. These syntactic formats and their examples in real-world texts are also included. In addition, KB@verb identifies nine major grammatical functions of verbs and their degrees of membership. Finally, the retrieval system and the hardcopy of KB@verb are illustrated.

Keywords: *The Syntactic-Semantic Knowledge-Base of Chinese Verbs* (KB@verb); sub-classes of verbs; semantic roles; syntactic formats; grammatical functions

0 引言

《汉语动词句法语义功能信息词典暨检索系统》(以下简称《动词句法语义信息词典》), 是袁毓林教授主持的《北京大学现代汉语实词句法语义功能信息词典》(简称《实词信息词典》)的一个子系统。本文主要介绍《动词句法语义信息词典》的知识内容和体系结构, 包括该词典为动词所划分的 8 种动词小类及其定义、为动词设置的 22 种语义角色及其定义、由目标动词与这些语义角色的不同配置而造成的 20 来种句法格式及其例句, 以及动词的 9 种主要

的语法功能及其对于该词类的隶属度。

1 《动词句法语义信息词典》的内容结构和理论背景

《动词句法语义信息词典》主要描写动词跟其他词语的搭配方式和语义上的句型构造。通过语料调查我们可以发现, 动词在句子中主要做谓语或谓语核心, 表示人或事物的某种动作、行为、状态或关系; 跟它搭配的词语表示这种动作、行为、状态或关系的主体, 或者受到这种动作、行为影响的事物, 以及工具、方式、处所等其他相关的事物。简单地说, 动词

收稿日期: 2021-02-21 定稿日期: 2021-03-29

基金项目: 国家科技创新 2030“新一代人工智能”重大项目(2020AAA0106701); 国家社会科学基金(18ZDA295)

的有关搭配成分是论元；论元又可以根据它们在动词所表示的事件中的作用(扮演的角色)分为不同的语义角色。

这样,了解一个动词跟什么样的语义角色搭配、它们怎样搭配、它们组合起来以后描述了一种什么样的情景,也就基本掌握了一个动词的意义和用法。难怪美国著名的认知心理学家 Steven Pinker 要说:“一个动词不仅仅是一个用于指称动作或状态的词,它实际上是句子的‘底盘’。它为该句子的其他成分提供了一个多槽的框架,在这里,无论主语、宾语还是各种介词宾语和从句等成分都可以各就各位,各司其职”^[1]。事实的确如此,请看例子:

例1 吐鲁番的葡萄熟了。——(主体+动词)

例2 萧何月下追韩信。——(施事+处所+动词+受事)

例3 我们向幸福进发。——(施事+目标+动词)

例4 黄鼠狼给鸡拜年。——(施事+对象+动词)

例5 我的婚姻我做主。——(范围+施事+动词)

例6 我家住在黄土高坡。——(主体+动词+处所)

例7 赵普用半部《论语》治天下。——(施事+工具+动词+受事)

例8 爹爹送我两尺红头绳。——(施事+动词+与事+受事)

从括号中对这些句子的简单的语义分析可见,意义和用法不同的动词能够跟不同的语义角色。通过不同的配置方式造成不同的句式,从而表达了不同的意义、描绘了不同的场景、讲述了不同的故事、报道了不同的事件。因此,可以直观地把动词想象成各种带钩子的原子,不同的动词所带的钩子的数量和形状不尽相同;于是,它们所能勾住的其他成分的数量和性质也就有所不同。如果借用化学上原子化合和配价的说法,那么动词就是语句组合的核心,像施事、主体、受事、范围、处所、工具等伴随成分就是配价成分。不同的动词有不同的配价功能,支配不同数量和不同性质的配价成分,构成不同形式的短语和句子。

《动词句法语义信息词典》就是从这样一种“情境语义学”(Situation Semantics)和“配价语法”(Valence Grammar)的角度,通过对大规模真实文本语料的调查和分析,来全面、准确、简明地描写动词在情境意义和搭配用法上的关键性特点,使读者“观其伴,会其意;明其价,知其用”;即让读者在查阅一个动词的条目以后,可以了解该动词通常跟哪些伴随成分一起出现,从而从搭配关系上理解该动词

的意义,明白其配价组合方面的特点,并掌握其基本的常用句式。

下面,介绍这部词典的动词小类、语义角色、句法格式和语法功能的描写体系。

2 《动词句法语义信息词典》的动词小类及其定义

这里的动词是根据其用法(即语法功能)来定义的,指具有下面这些语法功能的词:①可以作谓语或谓语核心;②可以受否定词“不、没有”修饰;③可以带真宾语或者不能受“很”等程度副词修饰。为了反映动词内部不同的词之间在有关用法上的差别,我们又把动词分为八种小类:<不及物动词>、<准及物动词>、<体宾动词>、<助动词>、<形式动词>、<名动词>、<强谓宾动词>、<弱谓宾动词>。

在本词典中,对于每一个动词词条,我们首先根据它们能否带真宾语而区分为“及物动词 vs.不及物动词”两大类:①可以带真宾语的是及物动词;②不能带真宾语是不及物动词。在广义的不及物动词中,那些能额外带有一个需要介词引导的必有性配价成分的是准及物动词。如果是准及物动词,那么直接标上<准及物动词>;剩下的是狭义的不及物动词,那么直接标上<不及物动词>。

如果是及物动词,我们根据它们能否带谓词性宾语而区分为“谓宾动词 vs.体宾动词”两大类:可以带谓词性成分作宾语的是谓宾动词,不能带谓词性成分作宾语的是体宾动词。如果是体宾动词,那么直接标上<体宾动词>。

如果是谓宾动词,那么还要细分为下列五种小类:强谓宾动词、弱谓宾动词、助动词、形式动词,名动词。粗略地说:只能带谓词性成分作宾语、表示情态意义的是助动词,直接标上<助动词>。只能带名动词、事件名词作宾语的是形式动词,直接标上<形式动词>。可以兼作名词的是名动词,直接标上<名动词>。剩下的谓宾动词,只能带谓词性成分作宾语的是强谓宾动词,标上<强谓宾动词>;既能带谓词性成分作宾语、又能带体词性成分作宾语的是弱谓宾动词,标上<弱谓宾动词>。

各种动词小类的具体定义和判定标准如下^[2]:

不及物动词:只能带一个必有性配价成分,不能带真宾语、可以带准宾语的动词。

准宾语包括下列三种:

(1) 量度宾语,即表示时量、动量或程度的宾

语。例如:

[休息了]一会儿 | [醒了]好几次
[锈了]一点儿

(2) 处所宾语,即表示运动的终点的宾语。

例如:

[来]北京 | [去]贵州 | [飞]乌鲁木齐
[到]亲戚家

(3) 存现宾语,即表示存在、出现或消失的宾语。

例如:

[来了]两个新队员 | [走了]几个客人
[新到了]一批冻猪肉

准及物动词:不能带真宾语、但是可以通过介词来引导与事、对象等必有性配价成分的动词。^[3]

准及物动词的配价成分,常见的包括下列五种:

(1) 与事,即表示动作行为的协作者。例如:

[跟]李红[结婚] | [和]对手[较量]
[与]浙江[接壤]

(2) 对象,即表示动作行为所针对的对象。

例如:

[向]朋友[求助] | [给]师傅[拜年]
[为]儿子[说情] | [怪罪][于]他人

(3) 施事,即表示动作行为的发出者。例如:

[由]王平[接手] | [由]老板娘[主事]
[由]我爸爸[做东]

(4) 处所,即表示运动的起点或终点等地点。

例如:

[起源][于]唐朝 | [出生][在]苏州
[毙命][于]海外 | [在]西安[落户]

(5) 依凭,即表示动作行为所依靠或凭借的事物。例如:

[以]音乐[见长] | [以]功臣[自居]
[以]失败[告终]

体宾动词:不能带谓词性宾语、可以带由体词性成分充当的真宾语的动词。

谓词包括动词、性质形容词、状态形容词和谓词性代词等。

体词包括名词、时间词、处所词、合成方位词、数量词、体词性代词等。例如:

骑(马) | 买(菜) | 捆(行李)
喝(一盅) | 驾驶(汽车)
修理(拖拉机) | 招待(他们俩)

强谓宾动词:只能带谓词性宾语、不能带体词性宾语的动词。例如:

打算(买房子) | 企图(越狱逃跑)

希望(大家都高兴)

觉得(这样做不妥当)

以为(没有人知道)

认为(老子天下第一)

弱谓宾动词:既可以带谓词性宾语又可以带体词性宾语的动词。例如:

看(下棋/立体电影)

喜欢(出去旅游/中式服装)

赞成(走出国门/他的意见)

反对(盲目地扩大产能/厂长提出的融资方案)

考虑(买一台电脑/这种办法)

助动词:只能带谓词性宾语、不能带体词性宾语的表示情态意义的动词。助动词又叫能愿动词、情态动词。例如:“能、能够、会、可以、可能、得、要、敢、想、应该、应当、该、愿意、情愿、乐意、肯、许、(不)配、值得”等。它跟不表示情态意义的其他谓宾动词的不同的语法特点是:①不能重叠;②不能带时体助词“着、了、过”;③可以放在“不~不”格式里。

形式动词:只能带名动词、事件名词或以它们为核心的体词性偏正结构为宾语的动词。形式动词又叫傀儡动词或准谓宾动词。下面是“形式动词+名动词”的例子:

有(准备/保障) | 作(调查/分析)
进行(处理/研究) | 加以(鉴别/管束)
给以(打击/报复) | 受到(批评/监控)
予以(表扬/追究)

下面是“形式动词+事件名词”的例子:

作(手术) | 进行(手术/战争)

名动词:可以作形式动词的宾语,可以修饰名词或受名词修饰的动词。名动词又可以称为名动兼类词。例如:

[有]准备/保障 | [作]调查/分析
[加以]鉴别/管束 | [给以]打击/报复
[受到]批评/监控 | [予以]表扬/追究

3 《动词句法语义信息词典》的语义角色系统及其定义

动词的各种伴随成分(即配价成分,又叫论元),根据它们跟动词在意义上的不同关系,可以区分为不同的语义角色。在本词典中,动词的配价成分首先分为必有论元和非必有论元两种,前者是构成意思相对完整的句子所不可缺少的,后者则用以扩充句子的意思,帮助形成意思相对复杂的句子。必有

论元可以分为主体论元和客体论元两种,前者主要作主语,后者主要作宾语。主体论元可以细化为施事、感事、经事、致事、主事等语义角色,客体论元可以细化为受事、与事、结果、对象、系事等语义角色。非必有论元可以从语义上分为依凭论元、环境论元和关涉论元三种,它们主要作状语。其中,依凭论元可以细化为工具、材料、方式、原因、目的等语义角色,环境论元可以细化为时间、处所、源点、终点、路径等语义角色,关涉论元可以细化为量幅、范围等语义角色。

上述 22 种语义角色的定义可以大致规定如下:

(1) 施事(agent, 简称为 A): 动词所表示的自主性动作行为的施行者。例如:

鸟儿飞了 医生来了
奶奶杀了一只鸡 爸爸买了一本书

在上面的例子中,“鸟儿”是发出“飞”这种动作的施事,“奶奶”是进行“杀(鸡)”这种行为的施事。

(2) 感事(sentient, 简称为 SE): 动词所表示的心理感觉、精神体验等事态的感知性的主体。例如:

我们看电影 | 爸爸认识李小龙
同事们赞成这个方案 | 大家讨厌这个人

在上面的例子中,“我们”是“看”这种感知行为的感事,“大家”是体验了“讨厌”这种心理行为的感事。

(3) 经事(experiencer, 简称为 EX): 动词所表示的某种变化的具有感知性的主体。例如:

同学们都毕业了 | 爷爷去世了
我碰到几个熟人 | 布什当选为美国总统

在上面的例子中,“同学们”是经历“毕业”这种变化性事件的经事,“我”是经历了“碰到”这种行为的经事。

(4) 致事(causer, 简称为 CAU): 某种致使性事件的引起者,即造成某种后果的致使性因素。例如:

大雨阻断了山区的交通
这一举措改变了公司的形象
罢餐事件引起了校方的注意
工厂倒闭的消息把大伙儿吓坏了

在上面的例子中,致事“大雨”是造成“山区的交通”发生“阻断”的引起因素,致事“罢餐事件”是“引起”“校方的注意”的致使性因素。

(5) 主事(theme, 简称为 TH): 动词所表示的性质、状态、关系或变化等事态的非感知性的主体。例如:

锅里的水开了 | 村后的石桥塌了
小孩掉水沟里了 | 我爸爸长了一个疖子

在上面的例子中,“锅里的水”是“开”这种状态的主事,“小孩”是“掉”这种变化的主事。

(6) 受事(patient, 简称为 P): 因施事或致事的行为而受到影响的事物。例如:

小猫逮耗子 爷爷喝葡萄酒
老师批评了王平 弟弟打碎了那面镜子

在上面的例子中,“耗子”是受到“逮”这种动作影响的受事,“王平”是遭受“批评”这种行为影响的受事。

(7) 与事(dative, 简称为 D): 动词所表示的动作、行为的非主动的参与者。例如:

老板对雇员发火
我们向当事人打听了一下

舅舅给了小明一本词典
同学们请教李老师一个地理方面的问题

在上面的例子中,与事“雇员”是“发火”这种事件的被动的参与者,与事“李老师”是“请教”这种事件的非主动的参与者。

(8) 结果(result, 简称 R): 由施事或致事的动作、行为造成的结果。例如:

爸爸又挖了一个菜窖
妈妈织了一件毛衣
猫咪在桌子上踩了一串脚印
弟弟把窗户纸捅了一个大窟窿

在上面的例子中,“一个菜窖”是“挖”这种动作的结果,“一串脚印”是“踩”这种行为的结果。

(9) 对象(target, 简称为 TA): 动词所表示的心理感觉、精神体验等感知性行为的对象和目标,有时可以用介词“对、对于”等引导。例如:

我爸爸认识吴校长 | 妹妹喜欢芭蕾舞
我奶奶居然相信通灵术
陈一平很熟悉广告业务 ~ 陈一平对广告业务很熟悉
我哥很热爱本职工作 ~ 我哥对本职工作很热爱

在上面的例子中,感事“我爸爸”所“认识”的对象是“吴校长”,感事“妹妹”所“喜欢”的对象是“芭蕾舞”。

(10) 系事(relative, 简称为 RE): 在动词所表示的状态、关系等事态里,跟主事、与事等相对的事物,一般表示相应于主事、与事等的属性、类型等。例如:

李四光是地质学家
 徐先生有两个女儿
 大家叫鞠萍知心姐姐
 通州属于北京

在上面的例子中,系事“地质学家”表示主事“李四光”的性质,系事“知心姐姐”表示与事“鞠萍”的称谓,系事“北京”表示主事“通州”的归属关系。

(11) 工具(instrument,简称 I): 动词所表示的动作、行为所凭借的器具,有时可以用介词“用、以”等引导。例如:

妈妈用水果刀切黄瓜
 王老师用显微镜看切片
 刘大夫用中药治风湿
 建筑师用计算机设计智能大楼

在上面的例子中,“水果刀”是“切”这种动作的工具,“计算机”是“设计”这种行为的工具。

(12) 材料(material,简称 MA): 动词所表示的动作、行为所用的材料,有时可以用介词“用、以”等引导。例如:

妈妈用毛线织了一双手套
 孙大爷用米泔水浇月季花
 小海娃用红蚯蚓引诱鱼群
 设计师用灯光装饰展厅

在上面的例子中,“毛线”是“织(手套)”这种行为所凭借的材料,“红蚯蚓”是“引诱(鱼群)”这种行为所用的材料。

(13) 方式(manner,简称 M): 动词表示的动作、行为所采取的方式、方法,有时可以用介词“用、以、经过、在(……下)”等引导。例如:

余子真用低音唱了一首民歌
 曼联队以点球取得了决赛胜利
 他们经过三年的奋斗编纂了一部词典
 电影节在社会各界的支持下顺利结束

在上面的例子中,“低音”是“唱(民歌)”这种行为的方式,“三年的奋斗”是“编纂(词典)”这种行为所用的方式。

(14) 原因(reason,简称为 RN): 动词所表示的动作、行为、事件等发生的原因,一般用介词“因、因为、由于、为、为了”等引导。例如:

鱼苗由于缺氧而死亡
 雷曼银行因金融危机而倒闭
 小明因考试失败而哭泣
 小芳为找不到工作而叹气

在上面的例子中,“缺氧”是造成主事“鱼苗”“死

亡”的原因,“找不到工作”是造成施事“小芳”“叹气”的原因。在表示原因的介词结构和动词之间经常用连词“而”来连接。

(15) 目的(aim,简称为 AI): 施事发出动词所表示的动作、行为、事件等的目的,一般用介词“为、为了”等引导。例如:

为了早日完工,工人们经常加班
 为了出国旅行,她兑换了一些美元
 为了加薪,小刚向厂长求情
 为了养家糊口,大伙儿拼命地工作

在上面的例子中,“早日完工”是驱使施事“工人们”“加班”的目的,“养家糊口”是促使施事“大伙儿”(拼命地)工作的目的。

(16) 时间(time,简称为 T): 动词所表示的动作、行为、事件等发生的时间,有时可以用介词“在、于”等引导。例如:

工人们在节假日也不休息
 我们明天去北京
 在上大学期间他们俩就相识了
1996年春天孩子出生了
 这件事发生在深夜

在上面的例子中,“节假日”是“(不)休息”这种行为发生的时间,“1996年春天”是“(孩子)出生”这种行为发生的时间。

(17) 处所(location,简称 L): 动词所表示的动作、行为、事件等发生的处所,有时可以用介词“在、于”等引导。例如:

同学们在教室里写毛笔字
 我们火车站会合
 在南操场上他们训练了一个上午
大礼堂里正在放映一部美国电影
抢劫案发生在银行外面

在上面的例子中,“教室里”是“写(毛笔字)”这种行为发生的处所,“银行外面”是“(抢劫案)发生”的处所。

(18) 源点(source,简称 SO): 动词所表示的动作、行为开始的地点或时间,有时可以用介词“自、从、在、于”等引导。例如:

赛车从北京出发
 我们从窗口看风景
 我们从七月初放假
 这些学员来自沿海地区
科举制度起源于隋朝
监狱里跑了一个犯人

在上面的例子中,“北京”是“出发”这种行为发生的源点,“七月初”是“(开始)放假”的源点。

(19) 终点(goal,简称GO):动词所表示的动作、行为结束的地点、时间或状态,有时可以用介词“朝、向、往、到、至、在、于”等引导。例如:

骏马向北方飞驰
我们朝窗外看星星
在南墙上他们挂了一幅山水画
黑板上老师又写了几行字
我们将工作到八月底
这些学员要分配到边远地区
科举制度延续至晚清
我们家里来了几个客人

在上面的例子中,“北方”是“飞驰”这种行为的终点,“八月底”是“工作”的终点。

(20) 路径(path,简称为PA):动词所表示的动作、行为、事件等中途经过的处所或时间,有时可以用介词“经过、通过、沿着、在、于”等引导。例如:

坦克从石拱桥上经过
火车经过天津抵达北京
士兵们沿着大马路巡逻
我从窗户通过两座楼房的空档仰望天空
我们将走四号线
他们通过了榆树林

在上面的例子中,“石拱桥”是“经过”这种行为的途径,“四号线”是“走”这种行为的途径。

(21) 量幅(extent,简称为EXT):动词所表示的动作、行为、事件等所涉及的数量、频率、幅度、时间长度等相关事项。例如:

小明去了两趟
鲁智深打了郑屠夫三拳
一个西瓜卖了八块钱
渔船偏离了主航道几百米
小王迟到了一刻钟
我们结婚二十多年了

在上面的例子中,“两趟”是“去”这种行为的量幅,“一刻钟”是“迟到”这种行为的量幅。

(22) 范围(range,简称为RA):动词所表示的动作、行为、事件等所涉及的具体方面,或者是主事所依附的主体;它是主体、客体、环境、依凭和量幅之外的事物,有时可以用介词“在(……上/方面)、关于、对于”等引导。例如:

在立法方面,全国人大做了许多工作
关于加快金融改革,刘建生教授提出了一个

方案

陈冬青在学业上取得了丰硕的成果
大楼施工,我们顺利地完成了第一阶段的工程

我的手表,发条断了
这本书,爸爸只读了第一章

在上面的例子中,范围“立法方面”是施事“全国人大”“做了许多工作”这种行为的具体方面,范围“我的手表”是主事“发条”所依附的主体。

4 《动词句法语义信息词典》的句法格式系统及其例句

在本词典中,动词跟其配价成分构成的句法格式,主要有三大类:第一类,是主体论元作主语、客体论元作宾语的“主语+动词(+宾语)”类基础句式;第二类,是非主体论元(包括:客体论元、环境论元、依凭论元、关涉论元)作话题的“话题+主语+动词(+宾语)”类派生句式;第三类,是用介词“把”引导客体论元,或用介词“被”引导主体论元的“把”字句或“被”字句等有标记的派生句式。例如:

例9 a. A+(用I+) +P:

妈妈正在洗衣服。|爸爸用大刀砍树枝。

b. A+(用MA+) +R:

奶奶炒了两个菜。|爷爷用柳条编了一个筐。

c. A+(在L+) +P/R:

哥哥在路旁种了槐树。|弟弟在墙上画了二幅画。

d. TH+ +RE(+EXT):

我姑姑是知识分子。|轮船偏离了主航道20多米。

例10 a. P+A+(用I+) :

衣服妈妈已经洗了。|那棵树爸爸用大刀砍了。

b. I+(A+) +P:

这块香皂洗衣服。|这把大刀爸爸砍树枝。

c. MA+(A+) +R:

这些香椿炒鸡蛋。|这些柳条我编一个箱子。

d. L+(A+) +P/R:

路旁种了槐树。|墙上弟弟画了一个太阳。

例11 a. A+(用I+) 把P+ :

妈妈把衣服都洗了。|爸爸用刀把那根树枝砍了。

b. P+被A+(用I+) :

衣服被妈妈都洗了。| 那棵树被爸爸用大
砍了。

c. A + 把 MA + ___ + R:

奶奶把肉末炒了两个菜。| 爷爷把柳条编了一个筐。

d. MA + 被 A + ___ + R:

肉末被奶奶炒了两个菜。| 柳条被爷爷编了一个筐。

上面的例子中,例 9 是一般性的“主语+介词结构+动词(+宾语)”类基础句式。其中,施事、主事等主体论元作主语,受事、结果等客体论元作宾语,工具、材料、处所等环境或依凭论元通过介词引导作状语。把这种状语放在圆括号中,表示没有它们也不影响句子的结构和意义的完整。例 10 是特殊性的“话题+主语+动词(+宾语)”类派生句式。其中,受事、结果等客体论元前置到句首作话题,其余部分作说明(用以对话题作出评论);施事、主事等主体论元作说明部分的主语,从而形成没有宾语的主谓谓语句。或者,工具、材料、处所等环境或依凭论元前置到句首作话题,其余部分作说明(用以对话题作出评论);施事、主事等主体论元作说明部分的主语,受事、结果等客体论元作宾语,从而形成保留宾语的主谓谓语句。例 11 是用介词“把”或“被”作标志的主谓句。在介词“把”的引导下,受事等客体论元可以前置到动词之前;在介词“被”的引导下,施事等主体论元可以降级为状语,同时受事等客体论元可以前置到句首作主语。

上面的句式表示,不用主语、宾语等句法成分概念,也不用名词(性成分)、动词(性成分)等句法范畴概念,而是用施事、受事、与事、工具、材料、时间、处所、范围、量幅等语义角色概念;目的是要反映动词的几个惯常性的伴随成分在语句中的出现位置,以及它们之间的相互共现和选择限制关系,从而帮助人们在纷繁多变的语句形式上发现语句的意义骨架的有限性。如果我们能够把不同的上下文抽象成有限的几种语义格式(即由语义角色作为分布框架的句法构式),那么就可以在语言学习中以有限统御无限了。

5 《动词句法语义信息词典》中关于动词的语法功能信息

在本词典中,我们根据袁毓林等(2009)和袁毓林(2010),设定动词的主要语法功能及其对于动词

这一词类的隶属度:

(1) 可以受否定副词“不”或“没有”修饰。符合得 10 分,不符合得 0 分。例如:

不走、不发展、没有前进、没有洗澡。

(2) 可以后附或中间插入时体助词“着、了、过”,或者可以进入“……了没有”格式。符合得 10 分,不符合得 0 分。例如:

躺着、买了(票)、去过(美国)、洗着澡、理了发、出过力、上过学;

钉子锈了没有、樱花谢了没有、尸体腐烂了没有、东京地震了没有。

(3) 可以带真宾语,或者通过“和、为、对、向、拿、于”等介词引导其必有论元。符合得 20 分,不符合得 0 分。例如:

看电影、踢足球、想心事、学习本领、研究太空、和朋友见面、为子孙造福;

对孩子发火、向观众挥手、拿次货充数、昆曲起源于昆山。

(4) 或者不能受程度副词“很”修饰,或者能同时受“很”修饰和带宾语。符合得 10 分,不符合得 -10 分。例如:

*很哭、*很做、*很改造、*很探索;

很想家、很怕考试、很感谢你们、很担心她的身体。

(5) 可以有“VV、V—V、V 了 V、V 不 V、V 了没有”等重叠和正反重叠形式。其中,V 代表动词。符合得 10 分,不符合得 0 分。例如:

坐坐、琢磨琢磨、瞧一瞧、试验一试验、说了说、调查了调查、吃不吃、考虑不考虑、饿了没有。

(6) 可以作谓语或谓语核心(因而一般可以受状语或补语修饰)。符合得 10 分,不符合得 -10 分。例如:

咱们走、钢丝断了、你们好好儿地玩、我们马上出发、修好、洗干净、整理出来、笑得前仰后合、跳得非常高。

(7) 不能作状语直接修饰动词性成分。符合得 10 分,不符合得 0 分。例如:

*站吃~站着吃、*笑说~笑着说。

(8) 可以跟在“怎么、怎样”之后,对动作的方式进行提问;或者可以跟在“这么、这样、那么、那样”之后,用以作出相应的回答。符合得 10 分,不符合得 0 分。例如:

怎么剪?~这么剪!、怎样开?~这样开!、怎么调查?~那么调查!、怎样发展?~那样发展!、双手

是怎么颤的?~这么颤的!植物是怎么发育的?~植物是这么发育的!

(9) 不能跟在“多”之后,对性质的程度进行提问;并且不能跟在“多么”之后,表示感叹。符合得10分,不符合得-10分。

6 《动词句法语义信息词典》的检索界面

图1是《动词句法语义信息词典》中检索系统的界面截图。

动词句法语义功能检索系统	
检索词 <input type="text"/> 检索	
词目	词目释义与句法语义功能
擦01	词目: 擦01
擦02	双语拼音: cā
擦03	词类属性: 体宾动词
擦04	词义解释: 用手或布、毛巾等摩擦物体使干净。跟“抹、拭”相近。
擦05	近义词: 抹、拭
近义词 擦	反义词:
近义词 抹01	其他形式:
近义词 擦02	近义词 抹02
语义角色:	施事A: 用手或布、毛巾等摩擦物体使干净的人; 受事P1: 施事所擦的物体,如“桌子”等; 受事P2: 被施事擦掉的東西,如“眼泪、鼻涕”等; 工具I: 擦拭时所使用的工具,如“抹布”等; 源点S0: 受事P2被擦掉前所在的地方,一般由受事P1转化而来; 终点G0: 受事P2被擦掉后所到的地方。
句法格式:	S1: A + (用I +) + P1/P2/S0 如: 服务员正在~桌子。 母亲在偷偷地~眼泪。 他~了~眼角上边。 小胖用手~了一下嘴。 他用手绢~眼泪。 他用手绢~了~眼角上边。 S2: P1/P2/S0 + (A +) + 了 如: 桌子~过了。 眼泪~了。 脸上的墨水印你快~了。 上头我~过了。 S3: P2 + (A +) + G0 如: 眼泪都~手绢上了。 手上的墨水他都~衣服上了。 S4: I + (A +) + P1/P2 如: 这块布(我)~桌子。 这条手绢他~鼻涕。 S5: A + 把P1/P2/G0 + 了 如: 妈妈刚把桌子~了。 你快把眼泪~了。 服务员把上面~了~。 S6: A + 把P2 + 从S0 + 去/掉 如: 小惠把泪水从眼角~去,头也不回地走了。 他把那些字从黑板上~掉了。 S7: P2 + 被(A) + 了/ + 掉/去 如: 泪水已被~去。 污渍被她~了。 车门上的锈迹被司机~掉了。 S8: A + 把P2 + 去 + G0 如: 他把眼泪都~手巾上了。 小孩儿把颜料~衣服上了。
语法功能:	1. 可以受否定副词“不”或“没有”修饰。如: 他从来不~眼镜片。 她今天还没有~饭桌。 2. 可以后附或中间插入时体助词“着、了、过”,或者可以进入“……了没有”格式。如: 她正~着杯子盖呢。 我~了一下鼻子。 那面镜子我~过了。 脸上的汗水~了没有? 3. 可以带真宾语,或者通过“和、为、对、向、拿、于”等介词引导其必有论元。如: 服务员正在~桌子。 母亲在偷偷地~眼泪。 4. 或者不能受程度副词“很”修饰,或者能同时受“很”修饰和带宾语。如: *很~ 5. 可以有“VV、V-V、V了V、V不V、V了没有”等重叠或正反重叠形式。其中V代表动词。如: 快~一~眼泪。 他~了~眼镜片。 桌子~了没有? 6. 可以作谓语或谓语核心(因而一般可以受状语或补语修饰)。如: 他正在~玻璃。 服务员正认真地~桌子。 眼镜片~得干干净净。 7. 不能做状语直接修饰动词性成分。 8. 可以跟在“怎么、怎样”之后,对动作的方式进行提问;或者可以跟在“这么、这样、那么、那样”之后,用以作出相应的回答。如: 你怎么~窗户的? 我这么~窗户的。 9. 不能跟在“多”之后,对性质的程度进行提问;并且不能跟在“多么”之后,表示感叹。如: *多~? *多么~! 隶属度: 100 查看隶属度量表
© 2015-2022 北京大学中国语言研究中心/中文系	

图1 《动词句法语义信息词典》检索系统的界面截图

7 关于纸质版《汉语动词造句词典》

我们把《动词句法语义信息词典》进行了简化处理,特别是删去了其中的“语法功能”这一板块,形成了纸质版本的《汉语动词造句词典》^①,以方便广大读者携带和使用。

参考文献

- [1] Pinker S. 思想本质: 语言是洞察人类天性之窗[M]. 张旭红,梅德明,译. 杭州: 浙江人民出版社,2015.
[2] 朱德熙. 语法讲义[M]. 北京: 商务印书馆,1982.

(下转第45页)

① 将由商务印书馆出版。

for symptom name recognition in free-text clinical records of traditional Chinese medicine: An empirical study[J]. Journal of Biomedical Informatics, 2014, 47: 91-104.

- [24] 杨锦锋, 关毅, 何彬, 等. 中文电子病历命名实体和实体关系语料库构建[J]. 软件学报, 2016, 27(11): 2725-2746.
- [25] 苏嘉, 何彬, 吴昊, 等. 基于中文电子病历的心血管疾病风险因素标注体系及语料库构建[J]. 自动化学报, 2019, 45(002): 420-426.



常洪阳(1996—), 硕士研究生, 主要研究领域为自然语言处理。
E-mail: hychang@fonmail.com



马玉团(1996—), 硕士研究生, 主要研究领域为自然语言处理。
E-mail: yutuanma@163.com

- [26] 鲁红英, 韩杨超, 范亚鑫, 等. 中文症状知识库的建立与分析[J]. 中文信息学报, 2020, v.34(04): 33-40.
- [27] Guan T, Zan H, Zhou X, et al. CMeIE: Construction and evaluation of Chinese medical information extraction dataset[C]//Proceedings of the CCF International Conference on Natural Language Processing and Chinese Computing. Springer, Cham, 2020: 270-282.
- [28] Artstein R, Poesio M. Inter-coder agreement for computational linguistics[J]. Computational Linguistics, 2008, 34(4): 555-596.



鲁红英(1966—), 博士, 教授, 主要研究领域为自然语言处理。
E-mail: iehyzan@zzu.edu.cn

(上接第 36 页)

- [3] 袁毓林. 汉语配价语法研究[M]. 北京: 商务印书馆, 2010.
- [4] 袁毓林, 马辉, 周韧等. 汉语词类划分手册[M]. 北京:

北京语言大学出版社, 2009.

- [5] 袁毓林. 汉语词类的认知研究和模糊划分[M]. 上海: 上海教育出版社, 2010.



袁毓林(1962—), 教授、博士生导师, 主要研究领域为理论语言学和汉语语言学, 特别是句法学、语义学、语用学、计算语言学和中文信息处理等。
E-mail: yuanyl@pku.edu.cn



曹宏(1968—), 副研究馆员, 主要研究领域为应用语言学和博物馆学。
E-mail: caohong@pku.edu.cn