

现代汉语述语动词框架研究

张庆旭

(中国人民大学语言所)

内容提要:述语动词框架是指述语动词同其周围名词或名词性成分的语义组合模式,本文在论述了述语动词框架研究的意义之后,着重研究了现代汉语述语动词框架的分类和述语动词框架内部的语义限制问题。

A Study on Predicate Verbs Frames of Contemporary Chinese

Zhang Qing-xu

Renmin University

Abstract: In the paper we give a category for the predicate verbs frames of contemporary Chinese and a discussion on the semantic restrictions within the frames.

一、研究述语动词框架的意义

述语动词框架是指述语动词同其周围名词或名词性成分的语义组合模式。人类社会已进入信息化时代。在信息化时代的语言研究要兼顾人、机两用,即要建立和发展人、机两用的语言学。(1)述语动词框架是人和计算机理解、运用语言的关键。述语动词框架研究,就是一个符合人、机两用需要的科研课题。语言学家吕叔湘先生指出:“动词为什么重要?因为在某种意义上,动词是句子的中心、核心、重心,别的成分都跟它挂钩,被它吸住”。(2)计算语言学家黄昌宁先生指出:“从计算机对自然语言的理解或翻译来说,述语动词和形容词是句子结构和语义解释的中心。因此,如果机器词典能对句子中的述语动词与其周围的名词性成分所发生的语义组合关系作出具体详尽的描写,就可以大大提高自然语言理解系统或机器翻译系统的性能”。(3)吕叔湘先生和黄昌宁先生分别从理论语言学的角度和计算语言学的角度论述了研究述语动词及其相联系的前前后后名词的重要意义,实际上就是论述了研究述语动词框架的重要意义。下面根据笔者的认识再从两个方面来说明述语动词框架研究的实际意义。

1. 有利于语言教学,包括语文教学和对外汉语教学。中、小学生的写作练习中,经常出现搭配不当的句子,如“雷锋的高尚品质浮现在我的脑海里”。这句话的述语动词“浮现”和它前面

* 本项研究由中国国家自然科学基金资助。

* 本文是从我的硕士学位论文《现代汉语谓语动词框架研究》中摘取出来的一部分修改而成的。在写作过程中得到我的导师林杏光教授的指导和帮助,谨此致谢。

——作者

的名词“品质”不搭配。为什么不搭配呢？我们可以从述语动词框架中的语义限制的角度来分析。“浮现”具有（+具体+形象）的语义特征，“品质”具有（-具体[-抽象]-形象）的语义特征，可见二者在语义制约上不协调，因此搭配不当。在外国人和我国兄弟民族学习汉语的过程中，词语搭配的问题更为突出。外国留学生学习汉语，在词类问题上存在的问题主要是动词和形容词，其次是名词和副词。而在动词上所犯的错误集中在动词跟后面名词的搭配以及有关“时、态”的运用上。（4）全面研究现代汉语述语动词框架对解决外国留学生在动词学习中的问题无疑是大有帮助的。而且运用述语动词框架来进行教学，可以用模式化的语言学分析代替泛泛而谈的语文评议，可以用语义特征分析和语义限制分析对词语搭配的问题进行具体的解释。

2. 有利于语言信息处理。随着信息化社会的发展，信息量急剧增加，而信息载体的85%以上都是自然语言。计算机对自然语言的理解和翻译是语言信息处理的重要技术。这种信息处理技术离不开机器词典。机器词典中需有述语动词框架的信息。有了这种信息，在语言分析过程中，当述语动词的义项确定以后，就可以获取正确的基于述语动词框架的句意表达；在语言生成过程中，根据给定的句意表达式和语用信息（如言谈焦点和言谈者的意图等），就可在机器词典中找到相应的述语动词框架及其论旨角色，进而正确生成出表层的句式。这些方法，在我国的一些大型的自然语言理解系统和机器翻译系统中已经得到验证。相反的，有的机器翻译系统，由于在词典中没有给出每个动词的框架信息，造成分析和生成频频出错，这便从反面验证了述语动词框架的使用价值（5）。

正是因为研究述语动词框架具有如此重要的意义，所以清华大学计算机科学系和中国人民大学语言文字研究所合作研究的科研项目《现代汉语述语动词机器词典的建立与研究》得到了国家自然科学基金的资助。

二、现代汉语述语动词框架的分类

分类是人类认识事物特征的一种手段，也是人类对事物认识的一种结果。因此，任何一门学科都有分类和归类的问题。对现代汉语述语动词框架进行分类和归类，可以帮助计算机掌握纷繁复杂的语言现象（6）。

对述语动词的框架进行统计、整理和分类是一项非常繁杂细致的工作，要弄到准确无误是相当不容易的。这项工作在林杏光老师指导下，我们先后进行了三次。第一次的统计结果写在我的硕士学位论文《现代汉语谓语句动词框架研究》上。第二次的统计结果发表在林杏光先生所写的《现代汉语动词大词典》的“使用说明”上。由于语料的限制，我们发现前两次的统计都有不够准确的地方，于是进行第三次统计。发表在这里的是第三次的统计结果，比前两次有进步，准确度有了提高。我们所依据的统计材料有三个：林杏光和鲁川两位先生组织王玲玲、孙德金等编纂的《动词大词典》和《现代汉语动词大词典》，清华大学计算机科学系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》。我把第三次统计、整理出来的述语动词框架分为3大类、9中类、58小类（7）。

一、一价述语动词框架：

1. 一价自动词框架：

（1）施事+V。如爆发、抱歉、挑动、奋斗等。

2. 一价内动词框架:

〈2〉当事+V。如爆、报销、出生、蒸发等。

二、二价述语动词框架:

3. 二价他动词框架:

〈3〉施事+V+受事:如败坏、踩、打倒、鼓吹等。

〈4〉施事+V+结果:如出版、变成、成立、修建等。

〈5〉施事+V+受事或结果。如雕刻、煎、刨、梳等。

〈6〉施事+V+受事或工具。如照2等。

〈7〉施事+V+受事或处所。如填2、堵、离开、照1等。

〈8〉施事+V+受事或系事。如演等。

〈9〉施事+V+受事或范围。如量、领导、考等。

〈10〉施事+V+受事或材料。如填1等。

〈11〉施事+V+受事或目的。如射1、淘1、拉1(6)等。

〈12〉施事+V+受事或与事。如指点、指导、追问等。

4. 二价自动词框架:

〈13〉施事+V+同事。如联合、联络、配合、认识等。

〈14〉施事+V+原因。如愁、躲、庆祝、躲避、算计等。

〈15〉施事+V+与事。如贷款1、屈服、祝福、祝贺等。

〈16〉施事+V+系事。如扮演、担任、当、化装等。

〈17〉施事+V+目的。如等候、交涉、竞选、追求1等。

〈18〉施事+V+依据。如吃2、遵循、遵照等。

〈19〉施事+V+工具。如玩儿2、压2等。

〈20〉施事+V+时间:如熬2等。

〈21〉施事+V+方式。如打19、打22、玩儿3等。

〈22〉施事+V+范围。如比赛、从事、讲究、争论等。

〈23〉施事+V+处所。如到达、接近、跨、住等。

〈24〉施事+工具+V。如撒气2等。

〈25〉施事+同事+V。如搏斗、寒暄、结合1、约会等。

〈26〉施事+与事+V。如报仇、道歉、送行、着想等。

〈27〉施事+原因+V。如操心等。

〈28〉处所+V+施事。如闪4等。

〈29〉施事+V+处所或时间:如过1、经过等。

〈30〉施事+V+处所或范围:如进入等。

〈31〉施事+V+目的或处所。如报考等。

〈32〉施事+V+与事或材料。如装备等。

5. 二价外动词框架:

〈33〉当事+V+客事。如包含、承受、懂、感染、笼罩等。

〈34〉当事+V+结果。如产生、变1、结2、开3等。

6. 二价内动词框架:

- 〈35〉当事+V+范围:如轰动、善于、擅长等。
- 〈36〉当事+V+工具。如挨(ai)2等。
- 〈37〉当事+V+数量。如差1、翻5等。
- 〈38〉当事+V+处所。如朝、达到、发源、闪现、在2等。
- 〈39〉当事+V+基准。如比5、超出等。
- 〈40〉当事+V+时间:如持续、度过等。
- 〈41〉处所+V+当事。如闹4(2)、出产、发生、没有2等。
- 〈42〉时间或处所+V+当事。如有2等。

7. 二价领属动词框架:

- 〈43〉领事+V+客事。如具备、具有、没有1、占有2等。
- 〈44〉领事+V+分事。如富有等。
- 〈45〉领事+V+客事或分事。如有1等。
- 〈46〉客事+V+领事。如属于等。
- 〈47〉分事+V+领事。如属等。

8. 二价系属动词框架:

- 〈48〉当事+V+系事。如代表2、等于1、顶7、姓等。

三、三价述语动词框架:

9. 三价他动词框架:

- 〈49〉施事+V+与事+受事。如给、颁发、告诉、贡献等。
- 〈50〉施事+与事+V+受事。如介绍1、载1(2)、捐赠等。
- 〈51〉施事+受事+V+与事。如传2等。
- 〈52〉施事+同事+V+结果。如攀2等。
- 〈53〉施事+材料+V+受事或结果。如换取等。
- 〈54〉施事+受事+V+材料或工具。如过3等。
- 〈55〉施事+V+受事+与事。如转达等。
- 〈56〉施事+V+受事+系事。如比4等。
- 〈57〉施事+V+受事+范围。如盘问等。
- 〈58〉施事+V+与事+结果。如安1等。

以上对述语动词框架所划分的3大类、9中类、58小类,反映了三个不同的分类层次,并且各个层次的分类标准不同。

3大类,是第一个层次。它的分类标准是动词的论元属性,即看动词能带几个必有论元。能带一个必有论元的动词是一价动词,由一价动词组成的框架即为一价述语动词框架;能带二个必有论元的动词是二价动词,由二价动词组成的框架即为二价述语动词框架;能带三个必有论元的动词是三价动词,由三价动词组成的框架即为三价述语动词框架。

9中类,是第二个层次,它是根据动词的6个次类划分出来的。动词的6个次类是:他动词、自动词、外动词、内动词、领属动词、系属动词。一价、二价、三价分别搭配动词的6个次类,本应是3乘6等于18类,但是一价中没有他动、外动、领属、系属,三价中没有自动、内动、外

动、领属、系属,所以去掉了9类只有9类。

58小类,是第三个层次,它的分类标准是动词论旨属性:指必有论元具体是什么论旨角色。论元属性指带几个必有论元,解决的是数量问题。论旨属性指所带的必有论元是什么论旨角色,解决的是质的问题。如“接、安、交涉、申请”这四个动词的必有论元都是两个,因此论元属性相同。但它们的论旨角色不同,前二个动词论旨角色是施事和处所,后二个动词的论旨角色是施事和目的,因而就构成二个不同的框架。

三、现代汉语述语动词框架的语义限制

述语动词框架的语义限制,应包含两个内容:一是述语动词对其前后名词的语义限制;二是前后名词对述语动词的语义限制。二个方面相互限制,而述语动词对其前后名词的语义限制是主要的方面。限于篇幅和时间,本文只描写述语动词对其前后名词的语义限制,即研究述语动词本身的性质对进入框架的名词在语义上有什么相应的要求,以避免生成出“草吃牛”、“月亮吃苹果”之类的错句。

清华大学计算机科学系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》描写了3008个述语动词框架,每个框架包含一个或几个论旨角色。本文所采用的研究方法是参照陈群秀先生编制的语义分类提纲,将各个论旨角色中具有代表性的名词进行语义类标注,然后对各个论旨角色的语义类分别进行统计,最后对3008个框架中全部论旨角色的语义进行统计,并将统计出来的语义类按一定的逻辑联系编制一个名词语义类表,以建立名词的动态分类体系。

3008个述语动词框架的全部论旨角色的语义类共212个,试按事、物、时空、部件的顺序排列如下:

一、1. 事

- 2. 事域 3. 交通运输 4. 科学技术 5. 学科学位
- 6. 事情 7. 事件 8. 事项

二、9. 物

- 10. 具体物
- 11. 生物
- 12. 人类 13. 人 14. 集体 15. 国家地区 16. 军事单位 17. 社会团体 18. 临时团体
- 19. 民族部落 20. 家庭家族 21. 机构单位 22. 人群 23. 新闻媒介
- 24. 动物 25. 兽类 26. 禽类 27. 鸟类 28. 家禽家畜
- 29. 植物 30. 农林作物 31. 粮食 32. 植物集合类名
- 33. 微生物
- 34. 非生物 35. 天然物 36. 天体物 37. 地表物 38. 地下物 39. 气象物 40. 自然灾害
- 41. 地形地貌物
- 42. 化合物 43. 陆地物 44. 水地物 45. 集合类地表物名
- 46. 人工物 47. 设施物 48. 建筑物 49. 器具物 50. 生产用具 51. 智能机器
- 52. 交通邮电工具 53. 交通工具 54. 空中交通工具 55. 陆上交通工具 56. 水上交通工具

57. 生活用具 58. 卧具
 59. 服装饰物用品 60. 服装 61. 娱乐用品 62. 文艺体育用品 63. 文艺用品
 64. 体育用品 65. 弹药
 66. 容器包装用具 67. 容器
 68. 原材料 69. 原料 70. 材料 71. 燃料
 72. 耗散物 73. 生物耗散物 74. 人类食物 75. 粮食饭菜 76. 药品 77. 嗜好品
 78. 饮料 79. 排泄物 80. 分泌物 81. 遗弃物
 82. 信息物 83. 作品 84. 文学艺术作品 85. 钱币 86. 抽象钱币 87. 票证
 88. 标志
 89. 人工物集合类名
 90. 现象物 91. 气象现象 92. 物理现象 93. 化学现象 94. 社会现象 95. 生理现象
 96. 病理现象
 97. 现象物的源和因 98. 生物力 99. 自然力 100. 能量
 101. 抽象 102. 精神活动工具 103. 声音 104. 符号 105. 信息 106. 自然语言
 107. 言语
 108. 精神活动的源和因 109. 责任 110. 权力 111. 理想 112. 感觉心境 113. 荣辱感
 114. 意向 115. 道德规范。
 116. 利益
 117. 活动和产物 118. 精神活动 119. 社交活动 120. 文体活动 121. 思维活动
 122. 精神活动产物
 123. 法律 124. 政策条令 125. 意见见解 127. 风俗习惯 128. 理论体系 129. 方式方法
 130. 规律 131. 知识
 132. 技术 133. 信念 134. 意志 135. 活动时机 136. 态势 137. 过程和历史
 138. 内容 139. 现象 140. 实质 141. 界限
 142. 差异 143. 障碍 144. 投入 145. 机能 146. 效验 147. 凭据 148. 前程
 149. 保障 150. 境界 151. 境遇 152. 风貌 153. 心理 154. 动作
 155. 行为
 156. 属性 157. 定量属性 158. 数 159. 数量 160. 比例 161. 速度 162. 空间属性
 163. 形状
 164. 重量 165. 质量 166. 气味 167. 构件 168. 价值 169. 形象 170. 生命
 171. 社会属性 172. 职称职务 173. 职业
 174. 心理属性 175. 程度 176. 关系 177. 原因 178. 来源 179. 条件 180. 目的
 三、181. 时空
 182. 时间 183. 具体时间 184. 有限时间 185. 时点 186. 时段 187. 节假日
 188. 节日纪念日 189. 空间 190. 具体空间 191. 自然界领域 192. 有限空间
 193. 范围 194. 人工范围
 195. 场所 196. 人工场所
 197. 抽象空间 198. 位置
 四、199. 部件

- 200. 具体物部件
- 201. 生物部件 202. 人的整体 203. 人体部件 204. 人体外部部件 205. 头部
- 206. 人体内部部件
- 207. 动物部件 208. 植物部件
- 209. 非生物构件 210. 器具物构件 211. 服装构件 212. 武器构件

林杏光先生认为,词语搭配是系统,词语搭配的系统表现为类。近 10 多年来,林先生先编义类词典,后编搭配词典,试图将义类词典中的语义类装进搭配框架,构成某一类词和某一类词的搭配,以建立汉语的词语搭配系统。结果是义类词典中的语义类不能装进搭配框架中去,究其原因是义类词典的语义分类属于静态分类,而且语义分类又有很大的灵活性。林先生总结了这经验,觉得研究语义聚合和语义组合的相互限制是认识语言结构和语言机制从而实现计算机模拟人类语言的一项重要研究,应该在组合框架中去研究聚合,即在动态框架中去归纳、统计语义类,进而建立动态的语义分类体系。我们这次正是采用了这种研究方法(8)。

迄今为止,国内外研究工作者建立义类体系(thesaurus)的方法,是对词义进行静态分类,他们并没有把语义放到一定的组合框架中去观察。所以他们建立起来的义类体系在自然语言处理系统中起的作用也是有限的。本项目的研究依据大规模语料,在 3008 个述语动词的组合框架中,以述语动词对有关论旨角色所施加的语义限制为标准,用动态的描写方法,来建立汉语名词性概念分类体系,这样的分类体系正是当前的自然语言处理系统的迫切需求。

注释:

- (1) 林杏光:《语言研究要注意人、机两用》(《语文建设》1993 年 2 期)。
- (2) 吕叔湘:《句型和动词学术讨论会开幕词》(见语文出版社 1987 年版《句型和动词》)。
- (3) 林杏光等主编《现代汉语动词大词典》序言,北京语言学院出版社 1994 年 11 月版。
- (4) 黄南松:《从对外汉语教学看现代汉语语法研究》(《语文研究》1993 年第三期)。
- (5) 林杏光:《使用说明》(见北京语言学院出版社 1994 年 11 月版《现代汉语动词大词典》)。
- (6) 林杏光:《基于格关系的现代汉语述语动词分类系统》(提交全国计算语言学第三届联合学术会议论文)。
- (7) 这里的述语动词框架是林杏光先生为首的课题小组整理的,参加整理的课题组成员是张庆旭、苗传江、李绘新、刘士娟、祁素芳等。
- (8) 林杏光:《论语义类和词语搭配》(中国人民大学学报 1991 年 5 期)。

主要参考书目

1. 黄昌宁:《序》(北京语言学院出版社 1994 年 11 月版《现代汉语动词大词典》)。
2. 《现代汉语动词大词典》的“使用说明”。
3. 林杏光:《进一步深入研究现代汉语的格关系》(陈力为主编《计算语言学研究与应用》,北京语言学院出版社 1993 年出版)。
4. 陈群秀:《汉语语义分类系统》(内部资料)