

# 汉语句型成分自动分析中谓语识别策略的研究

罗振声 孙长健 孙才

清华大学(100084)

**摘要:** 汉语句型成分自动分析的精度主要体现在对句子中谓语的分析。本文对汉语句型成分分析中谓语的识别进行了详细的讨论,并提出了通过语义关系进行校正的方法。

**关键字:** 句型成分、谓语、语义检查

## An Approach to the Recognition of Predicates in the Automatic Analysis of Chinese Sentence Patterns

Luo zhensheng Sun changjian Sun cai

Tsinghua University (100084)

### ABSTRACT

The accuracy of automatic analysis of the components in chinese sentence patterns is mainly reflected on the analysis of the predicates. This paper expounds the recognition strategy of the predicates in the analysis of chinese sentence patterns, and gives a method of rectification by semantics.

**Keywords:** Sentence Pattern Components, Predicate, Semantic Examination

### 一、引言

汉语句型的自动分析和分布统计是一项重要的基础性研究课题。汉语句型的综合频度统计,为现代汉语、基础汉语教学、自然语言处理以及机器翻译等领域提供了强有力的统计数据 and 科学根据。

以结构特征为标准的汉语句型系统中,从句型成分在句子中的次序和句型成分之间关系来看,谓语处于句子的中心和支配地位,主语、状语、宾语和补语的分析都是以谓语的 analysis 为基础的,因此,谓语的识别在整个分析过程中占有举足轻重的地位。谓语能否正确识别,直接关系到整个系统的分析精度。

鉴于上述原因,本文采用了以谓语为中心的句型成分分析方法:首先确定句子的谓语,然后根据谓语把句子分成两部分,对谓语前面的部分分析主语和状语,对谓语后面的部分分析宾语和补语。若句中出现多个谓语(假设 $n$ 个),并且这 $n$ 个谓语组成连谓式,则这 $n$ 个谓语把句子分成 $n+1$ 个部分,对第一个部分分析主语和状语,对第 $n$ 个部分分析宾语和补语,对其余 $n-1$ 个部分分析状语、宾语和补语。若句中出现兼语式或小句作宾语,则采用递归的方法分析该句子。

## 二、谓语的识别

### (一)、可作谓语的词

句子中作谓语的词主要是动词和形容词，体词和谓词性代词在某些情况下也能作谓语。下面将从动词、形容词、体词和谓词性代词作谓语的语法特点出发，讨论谓语的识别策略。

#### 一、动词性谓语

##### 1. 动词的细分类

本系统根据动词在句中的语法功能把动词分成以下十四个小类：

标注为：vg, vgo, vgn, vgv, vga, vgs, vgd, vgj, va, vc, vi, vy, vh, vv

##### 2. 动词性谓语的识别

由于动词经常作谓语，所以系统对动词的分析采用排除法，即对句中每个动词判断其是否满足不能作谓语的条件，若满足，则认定该动词不是谓语，否则，就认为该动词作谓语。系统在判断句中的动词是否作谓语时，主要依据以下两个方面：

##### ① 依靠该动词的小类标记来判断：

根据动词作谓语的可能性，可以把上述十四个小类动词分成如下三大类：

第一大类，标记为vg、vc、va的动词通常不作谓语；

第二大类，标记为vgo、vgs、vgd、vgj、vi、vy、vv的动词一般情况下作谓语；

第三大类，标记为vgn、vgv、vga、vh的动词可能作谓语，也可能构成谓词性短语作其它成分的修饰语。

##### ② 依靠该动词的上下文环境：

动词在句子中的位置比较灵活，除了能作谓语外，还可以构成谓词性短语作主语、宾语、定语、状语和补语等其它成分。但动词在充当这些成分时，还是有规律可循的，可通过以下上下文环境来判断：

##### 1). 助词“的”

谓词性短语作定语时和所修饰的中心词之间常用助词“的”来连接。如：

【例2-1】 履行合同 的 方式 很多。

┌───┐ | ┌┐

动宾结构+usde+中心词

##### 2). 数量词

根据汉语习惯，若有多个成分同时修饰一个中心词，数量词通常放在最前面。因此，可根据此规律将位于数量词与其所修饰的中心词之间作定语的谓词性短语识别出来。如：

【例2-2】 我们要做 一个 有理想、有文化、有道德、有纪律 的 好 学生。

┌┐ | ┌┐

┌──────────中间均为定语，可归并──────────┐

##### 3). 助词“地”

谓词性短语作状语时和所修饰的中心词之间常用助词“地”连接。“地”这种情况只限于由动词“有、没有、无”等组成的述宾结构和主宾同形的主谓结构。如:

【例2-3】我们 有条件 地 接受 了这个方案。

┌──┐ | ┌┐

述宾结构+usdi+谓语

#### 4).时间词

动词放在时间词前面可作时间状语。如:

【例2-4】解放 前 这里是有名的黄泛区。

┌───┐

#### 5).助词“得”

动词构成谓词性短语作补语时只能出现在句子末尾，且通常以助词“得”为形式标记。如:

【例2-5】她高兴 得 跳了起来 。

┌───┐

#### 6).谓语动词 vgv 和 vy:

动词构成谓词性结构作句子的宾语时，谓语动词只能是 vgv（带动宾）和 vy（动词“是”），若动词在谓语动词 vgv 或 vy 后，则可判断它作句子的宾语。如:

【例2-6】国务院 负责 管理全国税收工作。

┌┐ ┌──────────┐

vgv 谓词性宾语

#### 7).介词:

动词构成谓词性词组还可以作介词的宾语。作介词的宾语时，介词和动词之间除了修饰动词的副词性成分以外，不再有其它成分，因此，可通过动词前面的介词来判断该动词作该介词的宾语。如:

【例2-7】不要 把 遵守纪律 当成一句玩笑。

| ┌──┐

介词+谓词性结构

需要注意的是介词“被”和“给”（标注为 pbei）不满足上述原则，在这两个介词后面的动词是可以作谓语的。如:

【例2-8】他被打了一顿。

## 二、形容词性谓语

### 1.形容词的细分类

形容词一般分为性质形容词和状态形容词两类，标注分别为a和z。

### 2.形容词性谓语的识别

判断形容词作谓语和判断动词作谓语的策略正好相反，系统对句子中每个形容词判断其是否满足能作谓语的条件，若满足，则认定该形容词作谓语，否则不作谓语。

形容词作谓语有如下特点:

1).形容词往往在没有动词谓语的情况下才作谓语。

形容词作谓语时一般不带宾语，所以它作谓语的特点表现在补语的多样化。形容词作谓语时后面经常带“得”、“不”、介词结构、程度副词、复合趋向动词、数量词等补语和“了”、“的”等助词。另外形容词作谓语时，前面常有程度副词，否定副词等修饰。

2).形容词在有动词谓语的情况下也可以作谓语，但一定是动词在前，形容词在后。句式主要有以下两种情况：

(1).兼语式：

【例2-9】这个消息令人懊丧。

这种句式可以通过特殊动词 vgj 来确定。

(2).连谓式：

【例2-10】他听了很不愉快。

形容词在动词谓语后可能作谓语，也可以作补语。形容词作补语可分为结果补语、程度补语和可能补语。形容词作补语的特点是谓语和形容词之间不能有结构助词、宾语、趋向补语等成分，所以若谓语和形容词之间存在这些词，则可判该形容词作谓语。

### 三、体词性谓语

#### 1.体词的分类

体词分共八类：名词（nf,npf,npu,npr,ng）、时间词（t）、处所词（s）、方位词（f）、区别词（b）、数词（mx,mw,mg,mm,mh,mo,maf,mam,mab）、量词（qni,qnc,qnk,qng,qnm,qns,qnv,qnf,qnz,qt,qv）和体词性代词（m）。

#### 2.体词性谓语的识别

由于体词作谓语的情况比较少，所以系统判断体词作谓语的策略和形容词类似，都是直接判断该词是否满足其能作谓语的条件，若满足，则认定为谓语。

体词作谓语是受限制的：量词直接作谓语的情况极少，仅限于某些量词的重叠式。数词和数量词作谓语，多出现在说明日期、天气等句子里。体词作谓语的句式通常为：

（状语）+主语+（状语）+名词/名词词组/数量词组/时间词

【例2-11】今年小王已经十三岁了。

体词作谓语都是单谓语句，主语的前后经常有状语，可依靠这些特点来判断。

### 四、谓词性代词作谓语

谓词性代词是可代替动词或形容词的代词，标注为rp。如：“怎么样”、“这样”等。谓词性代词可以代替动词、形容词作谓语。如：

【例2-12】你怎么了？

谓词性代词作谓语的句子一般都比较简单，都是单谓语句。

### （二）、可作谓语的词组

常作谓语的词组有联合词组、偏正词组和主谓词组。

1. 联合词组作谓语通常为多个动词或多个形容词用“和”、“与”、“而”以及“、”等连接起来组成动词词组或形容词词组作谓语，因此，只需分别判断出这些动词和形容词作谓语即可判断出该句由联合词组作谓语。如：

【例2-13】我们应该 保持和发扬 自己的民族传统。

2. 偏正词组作谓语时其中心词基本上都是体词，因此，只需判断出偏正词组中的中心体词作谓语即可。如：

【例2-14】巴金 四川人。

3. 主谓词组充当谓语的句子称为主谓谓语句。主谓谓语句比较复杂，识别主谓谓语句的关键是识别出句子谓词前的成分包含两个主语。可以通过下列特征来识别主谓谓语：

(1). 主语中插入的一些表示修饰、限制、语气的成分。主要包括以下几类：

1) 逗号“，”，表示大主语与小主语之间的语气停顿：

【例2-15】外表文静的何士光，内心世界却非常丰富。

2) 语气词：

【例2-16】你这个人哪，脑袋可真灵。

3) 状语，主要包括副词、时间词、介词短语等：

【例2-17】小张在第一场比赛中脚就扭伤了。

(2). 主语中的代词。

可以通过谓语前的代词判断出谓语前由两部分组成，特别是代词在末尾时。如：

【例2-18】所有的房间他都打扫得干干净净。

(3). 处所主词

可以通过处所主语中常出现的方位词（如“里”、“外”、“上”、“下”等）来判断，可分为如下两种情况：

一种是大主语为处所主语，如：

【例2-19】院子里绿树成荫。

另一种是小主语为处所主语，如：

【例2-20】你心里委屈。

(4) 通过语义关系来判断。

主谓谓语句在语义上可分为以下几类：

1) 大主语和小主语之间存在领属关系，即大主语常常是一个整体或者一种范围，而小主语所指的或者属于大主语，或者是大语员的成员，或者处于大主语的范围里。例如：

【例2-21】我们学校 校园 大，学生 也多。

2) 大主语是受动者。例如：

【例2-22】这笔钱 谁 也不能动。

3) 大主语是谓语所涉及的对象。例如：

【例2-23】这个问题，我没有什么看法。

4) 小主语是受动者。例如：

【例2-24】他 一分钱 也不乱花。

