

# “VP + 的 + N”结构 N 省略的条件描述 \*

陆汝占 靳光瑾

(上海交通大学计算机系)

**摘 要:**偏正结构“VP+的+N”可分为两类,A类:中心词N可省略;B类:中心词不可省略。《现代汉语八百词》对分类条件有明确规定,本文对分类条件给出词汇义、句法结构、语义类型、逻辑表达式的形式描述。

## Description of Understood Construction “vp + de + (N)”

Lu Ruzhan Jin Guangjin

(Dept. of Computer Science Shanghai Jiao Tong University)

**ABSTRACT:** The constructions (vp + de + N) can be classified into the two classes: A. the central noun N is omitted (Understood) and B. others. In this paper the conditions on the classification are described.

汉语计算语义理论旨在汉语语义的形式化:量化、可操作化、可计算化,归结到一句话,就是要将汉语的语义分析定义为一种演算。自然语言的语义分析能够成为一种演算吗?英语的Montague语法和GPSG语法都已作到了这一点。汉语能否实现这一目标呢?这个实现过程有多大困难?现在只能说有一个初步的认识,即汉语的认知理论、模型完全不同于印欧语。分析方法上,后者在词汇、语法、语用、语境诸层面上有明显的界面区别,相互之间又有对应关系。这种情况可称为明显的分层性,简称为“面结构”。但是汉语则不同,各层面之间很难划分经纬,词法与句法之间没有明显的界限。句法、语义、语用这几个层面均与词的形、音、义有密切的关系,直接受到“字”这一因素的影响。词汇义贯穿各个层面,我们称为“体结构”。对于汉语的这种结构方式,如何进行句法、语义分析,又如何进行形式描述,是计算机处理汉语的中心问题之一。“的”字结构虽然是汉语中的一个局部结构,但它的分布非常普遍。(N1+的+N2)结构,笔者已作了一些讨论。本文是对“VP+的+N”结构中N省略的条件分析和形式描述。这看上去是句法分析问题,实际上单凭句法结构不能完成,要结合语义分析才能完成。偏正结构是现代汉语中最为普遍且最重要的语言结构,它集中地反映了汉语的特点。《现代汉语八百词》中规定:动+的+(名),中心词能作前面动词主语或宾语的,可省,否则不能。如动词已有宾语,只有当中心名词是表示动作的工具时才能省,否则不能。朱德熙先生在《“的”字结构和判断句》中根据动词的“向”和与动词\*词直接联系的NP之间的关系确定“VP+的”是否能指代“VP+的+N”。朱德熙先生认为当“N”处于潜主语或潜宾语的时候,“VP+的”可以指代“VP

\* 本文受国家自然科学基金资助。

+的+N”。

由动词性成分组成的“的”字结构修饰名词,同样存在两种不同的类,A类:“VP+的”可以代替“VP+的+N”;B类:“VP+的”不能代替整个偏正结构。本文选用例句,分析形成A类格式和B类格式的原因,每例均从下列四方面进行描述:

- 1)“VP+的+N”的句法格式,简记为符号JG;
- 2)N,VP的词汇意义,简记为符号CY;
- 3)N所属语义类型以及V所要求的语义类型,简记为符号YL;
- 4)“VP+的+N”的逻辑表达式,简记为符号LB。

对于例句先进行句法分析,给出句法格式(JG)。从句法格式中,可看出是否有(主、宾语位置上的)缺省(O),但究竟是主语位置还是宾语位置上,单从句法还不能判定,需借助于语义类型匹配。只有在缺省时,才能从句法关系上确定“的”后的中心词N是否有可能代入缺位。对于动词V及与之相关的NP作词义分析(CY),给出N的词汇义,以及语义类型(YL)、VP所要求搭配的NP的语义类型。至此,可给出句法结构的完整信息,最后译出语句和偏正结构的逻辑表达式(LB),提供了两类例句的比较依据。

文中花括号{}表示语义类型。本文所说的语义类型是借用数理逻辑及计算机科学中的“类型”(Type)一词。“类型”和“范畴”都是表示集合的概念。语义类型用一组特征集描述,简写也可用词语表示。用逻辑语言学的话来说,这组特征描述了一类或一种事物。论元、短语范畴组成一表达式,具有什么样的语义类型,就是指表达式所指称的对象是属于什么样的类、种。语义类型说明中规定了一个表达式及其相应的语义类型,写成如下形式,表达式:语义类型。采用“类型”这一概念的好处是可以运算,有类似于集合论中(元素)“属于”(集合)和“包含”(子集)的运算。“类型”还可以兼容,表示多个并集,称为多型类型。语言表达式的语法范畴说明规定了表达式属于什么范畴,表示为:表达式 $\in$ 范畴。

限于篇幅本文只讨论“VP相同N不同,N省略的条件”。关于“VP不同N相同,N省略的条件”以及N在其它环境下省略的条件,将另文讨论。

例:

A1 开车的人(开车的)

JG (O1+V+N2)+de+N

说明:句法结构是语句的句法分析结果。字母O表示缺省成分。

CY 开(二价,V[2]):发动或操纵。 说明:词汇义是基本设置。

YL N:{人},二价动词表示成 $V(x,y)$ (也可表示为谓词符号 $Kai(x,y)$ ),其中 $x:{人}$  $y$ :{车,船,机器,……},(多型类型) Gongju:{……}

说明:语义类型检查,N2:{车},由句法分析结果及N2的语义类型,可知N2与V[2]宾语的语义类型匹配,故N2为V[2]的宾语,主语空位,于是句法结构中V实际为V[2](O1,N2)。这一点不能单纯由句法层面给出,要依赖语义类型信息来完成。下同。

LB $\exists x,y(Ren(x)\wedge Che(y)\wedge Kai(x,y)\wedge\cdots X\cdots)$

说明:此处,Ren(x)表示x为人,Che(y)表示y为车,Kai(x,y)表示x开y。

1)逻辑表达式是就“VP+de+N”所在整个语句而言的,N指称某个个体在逻辑式中对应(个体)项,用论元符号x表示。这个论元符号x在逻辑式中什么位置上出现,决定于它在整个语句中是什么句法成分。所列例子如果仅是短语结构,它就没有办法确定其具体出现形式。

例如“开车的人走了”中 N 在句中作主语,相应的符号 x 作为一元谓词“走”的论元。语句“我看见开车的人”中 N(人)作宾语,相应的符号 x 作为二元谓词“看见”的(宾语)论元。)因此 x 在公式中只能用……表示它的地位和关系,但不外乎作为函数的论元(如称谓的领有者)、谓词的论元(主语论元、宾语论元)等。

2)存在量词以及相应的约束变元 x, y 谓词 Ren(人), Che(车), Kai(开), 作为附加条件—限制条件,正好对应 VP 修饰的 N。

3)如何从语言表达式得到这个逻辑表达式,通常的办法都是人手工完成的。(如《汉语逻辑概论》)。Montague 语法给出规则可自动实现这样的翻译。对于汉语的逐字逐句翻译(从句到公式)是汉语计算语义理论的内容,限于篇幅细节不列出了,只谈一点思路及方法。

根据上述的句法结构  $V[2](O1, N2)$  求解相应的逻辑表达式  $V[2](x, y)$ , 其中的变元 x, y 应该是约束的。 $V[2](x, y)$  及语义类型  $N: \{车\}$  与  $V[2]$  的第二论元 y 的语义类型匹配成功,意指“有车”且是“开”的对象,可得到:  $\exists y(Che(y) \wedge Kai(x, y))$ , 公式描述了 VP; 由空语类 O1, 语义类型  $N: \{人\}$  与  $V[2]$  的第一论元 x 的语义类型匹配成功,意指“有人”且是“开”的主体,可得到:  $\exists x(Ren(x) \wedge \exists y(Che(y) \wedge Kai(x, y) \wedge \dots x \dots))$ , 公式描述了  $VF + de + N$ 。

结论: N 可作  $V[2]$  的主语,可省,该例属于 A 类。

B1 开车的技术娴熟 (\* 开车的娴熟)

JG  $(O1 + VP + NP2) + de + N + (A)$

CY 开:(二价,  $V[2]$ ):发动或操纵技术:人类在利用自然和改造自然过程中积累起来并在生产劳动中体现出来的经验和知识,也泛指其他方面的技巧。技巧:表现在艺术、工艺、体育等方面的巧妙的技能。技能:掌握和运用专门技术的能力。

YL  $N: \{能力\}, V(x, y) x: \{人\} y: \{车, 船, 机器 \dots\}$

句法结构信息:  $V[2](O1, NP2) \wedge Gongju(O3)$

LB  $\exists P((\exists x, y(Ren(x) \wedge Che(Y) \wedge Kai(x, y)) \equiv P) \wedge F(P) \wedge Xian(F))$

意指: P 是命题,指“某人开某车”,其中  $F(p)$  表示 p 为技术,  $Xian(F)$  表示 F 为娴熟。

结论: N 与 V 所要求的语义类型不匹配,不能填入句法结构的空位, N 不能省,该例属于 B 类。

A2 游泳的孩子很多(游泳的很多)

JG  $(O1 + VP) + de + N + (A)$

CY 游泳(一价,  $V[1]$ ):人或动物在水里游动。

YL  $N: \{人\}, “孩子”是“人”的子类型; V(x) x: \{人, 动物\} (多型)$

句法结构信息:  $V[1](O1)$

LB  $\exists x(Haizi(x) \wedge Youyong(x) \wedge Duo(x))$

结论: N 可作 V 的主语,该例属于 A 类。N 在整个语句中又作主语且作 Duo 的论元。

B2 游泳的姿势很美 (\* 游泳的很美)

JG  $(O1 + VP) + de + N + (A)$

CY 游泳:同上例;姿势:身体呈现的样子。

YL  $N: \{样子\}, V(x) x: \{人\}$

句法结构信息:  $V[1](O1) \wedge Gongju(O2)$

LB  $\exists P(\exists x((Ren(x) \wedge V(x)) \equiv P) \wedge F(p) \wedge Mei(F))$ , 其中  $F(p)$  表示 p 为一种姿势

(样子), Mei(F)表示 F 美。

结论:N 与 V 所要求的语义类型不匹配,不能填入句法结构的空位,N 不能省,该例属于 B 类。

由 B1, B2 例可以看出, VP 表示的命题(一件事)本身就是 N 所指的事物类中的一件(开车也是一门技术,游泳也是一种姿势)。也就是说 VP 指称是 N 指称(集合)中的一个元素。此时 N 不能省。

A3 说话的人是个高个子(说话的是个高个子)

JG (O1 + VP + N2) + de + N + VP

CY 说(二价, V[2]):用话来表达意思。

YL N: {人}, V(x, y) x: {人} y: {话}

句法结构信息: V[2](O1, N2)

LB  $\exists x, y(\text{Ren}(x) \wedge \text{SHuohua}(x, y) \wedge \text{Hua}(y) \wedge \text{Gao}(x))$

结论:N 可作 V[2]的主语,可省,该例属于 A 类。

B3 说话的口音很熟(\*说话的很熟)

JG (O1 + VP + N2) + de + N + (A)

CY 说:同上例;口音:说话的声音

YL N: {声音}, V[2](x, y) x: {人} y: {话}

句法结构信息: V[2](O1, N2)

LB  $\exists x, y, z(\text{Ren}(x) \wedge \text{Shuohua}(x, y) \wedge \text{Hua}(y) \wedge \text{Kouyin}(z, \text{Shuohua}) \wedge \text{Shu}(z))$

结论:N 与 V 要求的语义类型不匹配,不能省,该例属于 B 类。

A4 我了解的情况不太多(我了解的不太多)

JG (NP1 + VP + O2) + de + N + (A)

CY 了解(二价, V[2]):打听、调查;情况:情形:事物呈现出来的样子

YL N: {情况}, V(x, y) x: {人} y: {情况}

句法结构信息: V[2](NP1, O2)

LB  $\exists x(\text{Liaojie}(\text{Co}, x) \wedge \text{Qingkuang}(x) \wedge \text{NOT Duo}(x))$

结论:N 可作 V 的宾语,可省,该例属于 A 类。

B4 我了解的目的只有一个(\*我了解的只有一个)

JG (NP1 + VP + O2) + de + N + (VP + NP)

CY 了解:同上例。目的:想要达到的地点或境地;想要得到的结果。

YL N: {地点, 境地, 结果……}, V(x, y) x: {人} y: {事实}

句法结构信息: V[2](NP1, O2) LB (略)

结论:N 与 V 所要求的语义类型不匹配,不能进入句法结构的空位,不能省,属于 B 类。

A5 游行的群众走过来了(游行的走过来了)

JG (O1 + VP) + de + N + (VP)

CY 游行:(一价, V[1]):广大群众为了庆祝、纪念、示威等在街上结队而行。队而行。

群众:泛指人民大众

YL N: {人}, V(x) x: {人}。

句法结构信息: V[1](O1)。LB(略)

结论:N可作V的主语,可省,该例属于A类。

B5 群众游行的路线已经确定了(\*群众游行的已经确定了)

JG (NP+VP)+de+N+(VP)

CY 游行:同上例; 路线:从一地到另一地所经过的道路

YL N:{道路},V(x) x:{人}。

句法结构信息:V[1](NP)。 LB(略)

结论:N与V所要求的语义类型不匹配,不可省,该例属于B类。

A6 扩大招生名额的学校有三所(扩大招生名额的有三所)

JG (O1+VP+NP2)+de+N+(有+NP)

CY 扩大(二价,V[2]):使(范围、规模等)比原来大。名额:人员的数额学校:专门进行教育的机构

YL N:{机构},V(x,y) x:{机构,……} y:{数额}

句法结构信息:V[2](O1, NP2)。 LB(略)

结论:N可作V的主语,可省,该句属于A类。

B6 扩大招生名额的问题很难解决(\*扩大招生名额的很难解决)

JG (O1+VP+NP2)+de+N+VP2

CY 扩大:同上句;名额:同上句;问题:需要研究讨论并加以解决的矛盾、疑难

YL N:{问题},V(x,y) x:{机构,……} y:{数额}

句法结构信息:V[2](O1, NP2)。 LB(略)

结论:N与V所要求的语义类型不匹配,不可省,该例属于B类。

A7.0 他说的话我没听清(他说的我没听见)

JG (NP1+VP+O2)de+N+(NP+VP)

CY 说(二价,V[2]):用话来表达意思。话的所指就是表达的意思。话:说出来的能够表达思想的声音,或者把这种声音记录下来的文字

YL N:{话},V(x,y) x:{人} y:{话}。句法结构信息:V[2](NP1, O2)

LB  $\exists y(\text{Ren}(C1) \wedge \text{Shuohua}(C1, y) \wedge \text{Hua}(y) \wedge \text{Tingqing}(C2, y))$

结论:N可作V的宾语,可省,该例属于A类。

A7.1 他说的办法可以试试(他说的可以试试)

JG ((NP1+VP+O2)+de+N+(VP)

CY 办法:处理事情或解决问题的方法。“办法”所指的是话的内容。

YL N:{方法}  $\cap$  = {话}, V(x,y) x:{人} y:{话}

句法结构信息:V[2](NP1, O2)。 LB(略) 结论:同上例。

B7 他说的时侯,我不在

JG (NP1+VP+O2)+de+N+VP

YL N:{时段},V同前。 句法结构信息:V[2](NP1, O1)。 LB(略)

结论:N与V所要求的类型不匹配,不能进入句法结构空格,不可省,该例属B类。

A8 他给我写的信还没收到(他给我写的还没收到)

JG (NP1+(P+NP)+VP+O2)+de+N+(VP)

CY 写(二价,V[2]):用笔在纸上或其他东西上做字 信:按照习惯的格式把要说的话

写下来给指定的对象看的東西

YL N: {书信}; V(x, y) x: {人} y: {字, 书, 信, ……}

句法结构信息: V[2](NF1, O2)。 LB(略)

结论: N可作V的宾语, 可省, 该例属A类。

B8 他给我写信的事儿你知道了(\*他给我写信的你知道了)

JG (NP1 + (P + NP) + VP + NP2) + de + N + (NP + VP)

CY 写: 同上句; 信: 同上句; 事儿: 事情

YL N: {事情}; V(x, y) x: {人} y: {字, 书, 信, ……}

句法结构信息: V[2](NP1, NP2)

LB  $\exists p(\exists x((Xie(C1, x) \wedge Gei(x, c2)) \Rightarrow P \wedge Zhidao(C3, p)))$

结论: 该句法格式中没有空位, N无法填入, 不可省, 该例属B类。

最后说明几点:

1. 句法结构信息是综合词汇义, 句法分析、语义类型匹配得到的。由此信息可进一步作语法语义分析。

2. 语义类型匹配在计算机科学中是很成熟的技术, 本文只说明一个思想方法, 没有介绍更多的运算的细节, 具体实现上可用合一算法。

3. 逻辑式的表示, 并不是唯一的, 只是为了能表达语句的涵义。如果有更丰富细致的逻辑系统, 描述能力将更强, 公式可以更精细一点, 但篇幅和范围都大大超出了。

4. 文中列出的四个方面, 是为了便于计算机实现而设立的。在A1, B1中有分别对应于短语VP及VP+de+N的逻辑式。它们在计算机处理中都是有用的。当人们说写汉语时, 凭语感知道什么样的短语结构可以省略名词, 什么样的短语结构不能省略, 实际上遵循了省略原则。反过来, 理解话语时, 也可凭语感知道什么样的短语结构实际上是省略结构, 而且能够补出省略的成分, 这也是根据省略原则。人们在生成话语和理解话语两个方向都是凭语感, 计算机处理时就应该有相仿的双向过程, 但必须基于理性分析和形式表示。当计算机生成汉语话\文时, 也可以将A类结构的短语省略名词, 操作上依据什么呢? 只要检查逻辑式中相应于名词的项变元的出现, 符合A类条件就行。在计算机理解时, 例如“开车的”, 应该理解成“开车的人”还是“开车的技术”(假定上下文中“人”“技术”都出现过), 哪一个合法的? 这时, 由“VP+de”的逻辑式 $\exists y(Che(y) \wedge Kai(x, y))$ 推导出来的逻辑式分别是例1中LB:  $\exists x, y(Ren(x) \wedge Che(y) \wedge Kai(x, y) \wedge \dots x \dots)$ 和例B1中LB:  $\exists p(\exists x, y(Ren(x) \wedge Che(y) \wedge Kai(x, y) \Rightarrow P) \wedge F(p) \wedge \dots)$ 前式中项变元x出现在谓词Kai的主语位上, A类。因此“开车的”理解成“开车的人”是合法的。后式中项变元F, 是命题变元, 它不出现在谓词Kai的主、宾语位上, B类。因此“开车的”不能理解成“开车的技术”。

## 参 考 文 献

- [1] 吕叔湘主编,《现代汉语八百词》,商务印书馆,1980
- [2] 朱德熙,说‘的’,中国语文,1961年第12期
- [3] 陈宗明,《汉语逻辑概论》,人民出版社,1993