

论元结构与句式变换*

詹卫东 北京大学

摘要 本文从句式变换的角度来审视动词的论元、论旨角色等语义概念，旨在说明：（一）“X元”动词跟“及物”动词与“不及物”动词等概念一样，是对动词的下位分类，不同类别的动词有不同的句式变换特征（或者说对不同句式的适应性不同），从这个意义上说X更适合看作是“符号值”，而不是“数值”，即便是把X看作是数值，它的取值范围也不应该仅限于0, 1, 2, 3这几个整数，而可以有1.5、2.5这样的小数。（二）“施事”“处所”等论旨角色概念也可以用来预示跟动词相关的句式变换性质，如果从句式变换等形式特征标准出发来认识“施事”等语义概念并相应地给出形式定义，可以更好地服务于计算机处理自然语言时的需要。

关键词 论元结构 论旨角色 动词 句式变换 形式规则 约束

一 引言

一般说来，“名词”“主语”等是所谓的句法概念，而“论元”“题元”“施事”等是所谓的语义概念。经过相当长时间的理论思考与实践，汉语学界对“名词”、“主语”这样的句法概念基本形成了共识，就是应该从功能分布角度或者说应该主要以形式标准来认识这些概念，但对于“论元”、“题元”、“施事”这样的语义概念，如何认识它们的性质，以及对这些概念应如何定义，目前似乎还存在不同的看法。徐烈炯、沈阳（1998）曾指出，“各种语言中对题元定义都会存在种种差异，解决这种问题，不能单单从某些动词与某些名词的直接词语联系来逐个判定，更需要找到一种整体的原则。”

我们曾经从计算机处理汉语时的实际需求的角度出发，讨论过语义范畴的确定原则以及语义范畴的相对性（詹卫东，2001）。本文拟进一步来讨论“论元”、“题元”等所谓的语义概念的实质，希望阐述这样两个观点：（1）就计算机处理自然语言的要求来说，对这些语义概念，要跟对待句法概念一样，研究如何以形式标准来加以定义，因为只有以形式标准加以定义，对计算机处理自然语言才有用，而且好用；（2）要以形式标准对“论元”、“论旨角色”等语义概念进行定义，一个主要的办法就是进行句式之间的变换测试，观察动词及其相关名词性成分对不同句式的适应性。

我们认为，在以计算机为背景来认识“论元”、“题元”、“施事”等概念时，上述两点应该被看作是技术层面上定义语义范畴的“整体原则”。

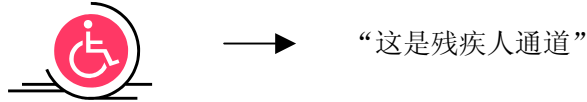
为了更好地阐述我们的看法，不妨先扯远一点，让我们从“什么是意义”这个古老的问题谈起。哲学和语义学对这个问题从不同角度给出了许多不同的理论表述（徐烈炯，1995）。虽然限于篇幅，本文不可能对这个问题加以展开，但不妨来看看，对于计算机来说所谓“知道一个语言形式的意义”是“什么意思”。我们的认识是：

（1）理解“意思”的过程，可以表示为对符号进行“变换”的过程

这样的符号变换可以在一种语言内进行，可以在不同语言之间进行，甚至也可以在不同性质的符号系统之间进行。比如：

* 本文研究工作得到“高等学校全国优秀博士学位论文作者专项资金”（项目编号200110）和国家973课题“面向中文信息处理的现代汉语动词论旨结构系统和汉语词语语义分类层级系统研究”（项目号：G1998030507-1）资助。本文初稿得到陆俭明先生的指教，特此致谢。文中错谬概由作者个人负责。

- a. 在一种语言内进行的符号变换
“张三打了李四” → “李四被张三打了”
- b. 在不同语言之间进行的符号变换
“张三用手打了李四” → “Zhang San hit Li Si with his hand”
- c. 在不同性质的符号系统之间（比如视觉符号系统与听觉符号系统）进行的符号变换



以上各例中，箭头左边的“表达形式”的“意义”，均可以用箭头右边的“形式”来表述。换言之，如果把箭头左边的形式作为输入，箭头右边的形式作为输出，输出端就是输入端的“意思”，而输入输出的过程，就是对输入符号串进行变换的过程。

(2) 这个符号变换过程没有止境¹

通过符号变换操作来表示一个语言成分（形式）的“意思”，有多种可能性。对此也可看作一个语言形式有一条很长的“意思”链。在日常语言使用中，人们在这个链上选取了一个当时（即确定的语境下）比较好用（有效）的环节，作为通常所说的“意思”。比如：

- d. “张三买了辆车” → 意思 1：张三把车买回来了
意思 2：张三拥有了一辆新车
意思 3：张三可以使用这辆车
意思 4：张三的钱都花完了

.....

如果把上面 a—d 例中左边的表达形式作为“输入”给一台计算机，计算机能够“变换”出 a—d 例中右边的那些形式作为“输出”反馈给人看，那么，人们很可能会惊呼“这台计算机懂我的意思了”。其实，人与人之间的交际过程也是在这样的输入输出的交替变换中进行的，正是在这样的变换中，人们相互之间就“显得知道对方的意思”了。

根据上面的描述，如果能够从一种形式 A 向另一种形式 B 做“正确的”（或“恰当的”）变换，B 就是 A 的一个“意思”。而为了保证“变换”的“正确性”，就要保证：

- (1) B 本身必须是一个“合语法”的形式；
- (2) B 跟 A 之间要满足某种“关系”（比如同义、蕴涵等等逻辑关系）。

实际上，人们设计出的句法范畴也好，语义范畴也好，目的都是为了保证这两个层面上的“正确性”²。引入“论元”、“施事”这些语义概念，跟“名词”、“主语”等句法概念一样，都是为进行正确的变换提供有效的条件。基于此，如果自觉地直接从“变换”的角度去认识和定义这些所谓的语义范畴，就能帮助计算机更好地去控制和模拟理解自然语言意思的过程。

二 动词的论元性质与句式变换

一般说来，动词的论元结构包括（1）动词的论元性质；（2）动词的论旨角色性质；（3）

¹ 或者说人们总是根据“需要”来停止变换操作，而“需要”是无止尽的，也是难以确定的。

² 从保证“变换的正确性”这个角度来说，语法研究的目的是为了保证最基本的正确性。人能够自由地进行理论上无止境的“变换”游戏，除了语法知识外，还需要逻辑知识，并且还涉及到历史文化、社会生活等百科知识，后者更是计算机目前难以掌握的，因而也不可能像人一样自如地进行“形式变换”。从这点上来说，“变换”显然是有层级的，“变换”的能力依赖于知识的多少。本文讨论的“句式变换”仅限于语法研究的范围内，只涉及语法条件制约的变换。暂不涉及到逻辑层面（比如“蕴涵”关系），更不涉及由百科知识控制的变换。

因此例 8 中 A 组和 C 组合法，B 组不合法。

例 9 中的动词（“发火”类）跟“顶嘴”类动词类似，也允许只有一个 NP（“主体”角色）与之同现。如果是与两个 NP 同现，其中一个 NP（“客体”角色）允许插入到双音节动词的两个音节之间，形成所谓的动词的“离合”用法。

例 10 中的动词（“撤职”类）跟“发火”类动词很类似，不同之处在于，如果是只有一个 NP 在动词前位置与动词同现，则与“撤职”类动词同现的是充任“客体”角色的 NP，而不是充任“主体”角色的 NP。

例 7—10 中 B 组都是不合法的表达形式，这是这些动词跟典型的二元动词的最大不同之处，因为典型的二元动词都可以在 B 组所代表的句式中出现，构成合法的表达形式。

下面例中的动词从意念上讲都涉及到三个 NP，但在句式变换上也有不同特征。

- (11) A B C
张三问了两个问题 —— 张三问了李四 —— 张三问了李四两个问题
张三教英语 —— 张三教李四 —— 张三教李四英语
- (12) A B C
张三拿了两本书 —— * 张三拿了李四 —— 张三拿了李四两本书
张三吃了两个苹果 —— * 张三吃了李四 —— 张三吃了李四两个苹果
- (13) A C D
张三发出了指令 —— * 张三发出了李四指令 —— 张三向李四发出了指令
张三打听消息 —— * 张三打听李四消息 —— 张三向李四打听消息

例 11—13 中的动词都可以跟三个 NP 发生语义关联，其中两个是指人 NP，一个指物 NP。例 11 中的动词（“问”类）在跟 NP 共现时可以有三种方式，分别是 A、B、C 所代表的三个句式；例 12 中的动词（“拿”类）在跟 NP 共现时有两种方式（A、C），不能以 B 的方式共现；例 13 中的动词（“发出”类）既不能以 B 的方式（限于篇幅，上面对比举例时未列出），也不能以 C 的方式与 NP 共现，而是以 A、D 两种方式与 NP 共现。方式 D 需要通过介词引出一个 NP 成分。

显然，从句式变换特征上看，“拿”类动词和“发出”类动词也跟典型的三元动词有不同。换句话说，这两类动词是否归属三元动词，也是一个问题。

透过上述几组句式变换的例子，不难看出，动词的论元性质实际上包含两个方面：论元的数量和论元的句法位置。仍用挖坑的比喻来说，就是一个动词能挖几个坑？以及动词把坑挖在哪儿？如果仅从名称上看，“一元动词”“二元动词”这些名称很容易让人只注意到前一个方面而或多或少地忽视后一个方面。但实际上这两个方面却是密不可分的，必须同时考虑这两个方面，才能对动词的论元性质有正确的认识。不同动词所能联系的论元 NP 的个数有多有少，NP 的位置相对于动词有前有后，有些动词的论元 NP 出现时还需要介词来帮忙。将动词与名词组配成句时，必须按照动词这两个方面的要求来组织句式，违反了其中任何一个方面的要求，都会造成错误的表达形式。由此，我们有理由对“一元”“二元”“三元”这样的“论元数”范畴提出下面两个问题，做进一步的反思。

(1) 所谓“论元数”，是“数值”，还是“符号值”⁷？

仅从字面上，人们很容易把“一元”“二元”等术语理解为“数值”，是在描述动词所能组配的 NP 的“个数”。但上文的句式变换实例已经清楚地显示，动词除了联系的 NP 个数有差异外，NP 所能占据的句法位置也有差异。如果把“一元”“二元”中的“一”“二”理

⁷ “符号值”的提法借用了计算机科学中“数据类型”(type of data)的概念。简单地说，“1”可以作为一个“符号”（“记号”），也可以理解为一个“数”，在北大的邮政编码“100871”中，“1”是“符号”，在“1+2=3”这个算术表达式中，“1”代表了“数值”。我们认为，在“1元动词”中，“1”更适宜理解为“符号”。在更抽象的层面上说，“符号”的作用在于“区别”，“数值”的作用在于“计算”。

解为“数值”，实际上是不利于说明句法位置上的差异的。事实上，“一元”“二元”就如同“及物”“不及物”一样，是对动词的进一步分类。这里的“一”“二”更适理解解为“符号值”，而不是“数值”。换言之，所谓的“论元数”与其强调其数目字上的意义，还不如看做是一个动词小类的名称更好⁸。

(2) 即便把“论元数”看作是“数值”，取值范围仅仅是 0, 1, 2, 3 吗⁹？

上文已经说明，动词的“论元”包含了动词所联系的 NP 的个数以及 NP 所能出现的位置两个方面，因而仅仅将论元的取值范围限制在 0—3 的整数范围内，就难以覆盖动词的各种类型，出现例 5 以及例 7—10，例 12—13 等不容易划入典型的一、二、三元动词集合中的情况。造成这种归属困难的部分原因大概是人们预设了动词“论元”的取值范围¹⁰。但事实上，论元的取值范围完全是可以拓宽的。可以从整数向小数拓宽，比如增设 1.5 价动词来对应“死”类动词，增设 2.5 价动词来对应“发出”类动词，等等，也可以是扩大整数范围，比如鲁川等（2000）就提出了“买”类动词属于“交易类四价动词”。

不难看出，增加“1.5 元动词”也好，提出“4 价动词”也好，都是在动词这个大类下分出一些个有自己句法特点的小类而已。如果不把这里的“1.5 元”和“4 价”等理解为数值，而只是看作是一个“类名”，那么不但可以避免一些难题和争议，而且可以把研究的注意力集中到挖掘一类类动词的句法特征上。说到底，研究动词的“论元”性质，就是从动名组配的角度来对动词进行分类，这跟以句法功能（分布）为标准区分出动词、名词、形容词等“词类”完全一样，不同小类的动词在句式变换上有不同的行为特征，给一个动词贴上“1 元动词”的标签，就意味着该动词不能像上文例 4 所显示的那样进行句式变换，而如果给一个动词贴上“1.5 元动词”的标签，就意味着该动词可以像上文例 5 显示的那样进行相应的句式变换。诸如此类。

基于上述分析，不难看到，如果能够在系统的句式变换基础上考察动词的论元性质（即从动名组配能力的角度区分出动词的不同小类），在计算机进行自动句法分析和句子生成的时候，就可以将动词的不同小类作为判断条件去控制不同的句式变换（规则示例见第四节）。

三 动词的论旨角色与句式变换

上一小节已经说明，不同小类的动词能够结合的 NP 数量不同，NP 所能出现的句法位置也有所不同。这是由动词的论元性质来刻划的。本小节将进一步说明，不同小类的动词能够结合的 NP 的类型也有差异，这是通过动词的论旨角色性质来刻画的。这里不妨继续用动词挖坑这个比喻来说明不同类动词的论旨角色性质差异：不同类动词能够挖出的坑的形状不同，有的动词可以挖出圆形和方形的坑，有的则挖出五角形或者三角形的坑。动词挖出的坑形状不同，在句式变换上也就会有相应的形式表现。正是依靠这些句式变换上的形式表征，我们才可以深入了解动词能挖出什么形状的坑，或者说，可以根据句式变换标准来定义动词的不同论旨角色。不过，由于研究工作还不够深入，目前还难以系统地以句式变换特征来定

⁸ 袁毓林（1998）提出“一元二位动词”“准二元动词”“准三元动词”等，实际上也可看作是在分类的意义上使用这些“数目字”。另外值得一提的是，朱德熙先生当年提出动词的“向”（即“论元”、“价”）的概念时，曾经用动词的“向数”来计算“VP 的”的歧义指数，但就无歧义来说，其实是并不需要数值意义上的“计算”的，根据动词和名词的条件判定即可，因而动词的“向”更主要的也是分类意义。

⁹ 也有的学者不同意 0 元动词的提法，这样的话，取值范围就仅限于 1, 2, 3 了。

¹⁰ 这很容易让人联想到句法研究中许多类似的产生过争议的问题：比如“看一下”中的“一下”是宾语，还是“补语”？“现在开会”中的“现在”是主语，还是“状语”？“张三吃了他三个苹果”是双宾语结构，还是单宾语结构？“春天的到来”是 np，还是 vp？等等。这些问题都是以选择问的方式提出，于是在给答案之前都好像已经给定了备选答案，即设定了问题解的取值范围。但这类有争议的问题实际上都是分类问题，研究者完全也可以在已有的分类之外，增加新的分类，即走“第三条路”给出答案，而不一定要拘泥于原有的分类（当然，我们也不反对在原先设定的分类框架下对这些问题进行深入探究）。

义动词的各种论旨角色，但即便是下面列举的一些情况，也足以阐明我们的观点，即句式变换特征可以而且应该用来作为像“施事”“受事”“工具”“处所”等论旨角色概念的形式定义标准。

下面先从陈述形式与疑问形式之间的句式变换来看动词论旨角色对不同句式的适应性差异。请看例子：

- | | | | |
|------|-------------|-------|----------------|
| (14) | A | | B |
| | 吃苹果 → 吃什么 | | 睡沙发 → 睡哪儿 |
| | 放钱 → 放什么 | | 放桌上 → 放哪儿 |
| | 晾衣服 → 晾什么 | | 晾铁丝上 → 晾哪儿 |
| | 写小说 → 写什么 | | 写黑板上 → 写哪儿 |
| (15) | 吃食堂 → * 吃什么 | * 吃哪儿 | (“食堂”作为“方式”角色) |
| | 吃大碗 → * 吃什么 | | (“大碗”作为“工具”角色) |
| | 洗凉水 → * 洗什么 | | (“凉水”作为“工具”角色) |
| | 写毛笔 → * 写什么 | | (“毛笔”作为“工具”角色) |

如果撇开动词不论，只看上面例 14—15 中动词后位置的名词，它们应该都可以自由地替换为相应的疑问代词形式（“什么”：问“事物”，“哪儿”：问“处所”），但如果联系这些名词跟动词形成的不同语义关系来看它们跟疑问形式的对应，就可看出一定的差异了。宽泛地说，出现在动词后位置，跟动词发生一定的语义关联的 NP，都可以笼统地称为动词的“客体”角色，但细究起来，这些“客体”NP 还可以区分为不同的类型，实际上例 14—15 中陈述形式与疑问形式之间对应关系的差别就显现了动词后“客体”NP 的不同情况。充任“受事”“结果”等语义角色的 NP 可以有“V 什么”这样的疑问形式（例 14A），充任“处所”角色的名词可以有“V 哪儿”这样的疑问形式（例 14B），而其他的充任“方式”“工具”等语义角色的名词则没有对应的“V 什么”“V 哪儿”这样的疑问形式（例 15）。

上述句式变换差异就可以用来作为“受事/结果”跟“工具/方式”论旨角色的一个形式判别标准，或者说作为“受事”等论旨角色的形式定义特征之一来更为严格地描述动词周围的 NP。比如例 15“吃食堂”中的“食堂”并不适合看作是“吃”的“处所”角色（参见王占华，2000），如果仅从意义上进行辨析，不容易说清楚，而如果像上面这样，从句式变换的形式特征上来认识和定义“处所”角色的话，就更清晰一些。

下面的几组句式变换的例子将进一步把“处所”角色做更进一步的区分。

在动词周围的“处所”NP 中，实际上存在不同情形，有的是表示整个 VP（对应着“事件”）所在的“空间”的，不妨叫做“事件处所”角色；有的则是表示参与事件的某个实体（包括“人”或“物”）所在的“场所”的，不妨叫做“实体处所”角色。通过下面的句式变换例子，我们将看到，“事件处所”与“实体处所”相对于动词有不同的位置要求。

- | | | | | | |
|------|-------------------------|----|------------|----|---------------|
| (16) | A | | B | | C |
| | * 在窗外下雨 | —— | 窗外下着雨 | —— | * 把雨下在窗外 |
| | * (张三) 在脸上流汗 | —— | (张三) 脸上流着汗 | —— | * (张三) 把汗流在脸上 |
| | * 在门口站人 | —— | 门口站着人 | —— | * 把人站在门口 |
| | * 在台上坐主席团 | —— | 台上坐着主席团 | —— | * 把主席团坐在台上 |
| (17) | * 在地上削土豆皮 ¹¹ | —— | * 地上削着土豆皮 | —— | 厨师把土豆皮削在地上 |
| | * 在地上剥香蕉皮 | —— | * 地上剥着香蕉皮 | —— | 张三把香蕉皮剥在地上 |
| | * 在地上推倒李四 | —— | * 地上推倒着李四 | —— | 张三把李四推倒在地上 |
| (18) | 在黑板上写字 | —— | 黑板上写着字 | —— | 老师把字写在黑板上 |
| | 在铁丝上晾衣服 | —— | 铁丝上晾着衣服 | —— | 她把衣服晾在铁丝上 |

¹¹ “在地上削土豆皮”本身是合语法的，但它不是“把土豆皮削在地上”的同义句式变换，例 17 均如此。

- 在墙上贴报纸 —— 墙上贴着报纸 —— 他把报纸贴在墙上
- (19) 在家里洗衣服 —— * 家里洗着衣服 —— * 他把衣服洗在家里
- 在脸盘里洗衣服 —— * 脸盘里洗着衣服 —— * 他把衣服洗在脸盘里
- 在食堂里吃饭 —— * 食堂里吃着饭 —— * 他把饭吃在食堂里
- 在图书馆找书 —— * 图书馆找着书 —— * 他把书找在图书馆
- 在书里找答案 —— * 书里找着答案 —— * 他把答案找在书里

上面 4 组例子包含了三种句法格式（这里用“NPL”代表“处所类 NP”，用“NLNP”代表“非处所类 NP”。为简单起见，C 格式中“把”前 NP 略去）：

A 组： 在 + NPL + V + NLNP

B 组： NPL + V + 着 + NLNP

C 组： 把 + NLNP + V + 在 + NPL

A、B 两格式中“处所类 NP”均出现在动词前，C 格式中“处所类 NP”在动词后位置。例 16—19 中的动词可以根据对上述三个句式的适应性不同而分为四类，下表描述了这四小类动词的特征：

动词小类	能否进入 A 格式	能否进入 B 格式	能否进入 C 格式	特征描述
I (“站”类)	-	+	-	(1) 动作所表达的事件在 NPL 发生； (2) NLNP 的位置也在 NPL。
II (“削”类)	-	-	+	(1) 动作所表达的事件并不在 NPL 发生； (2) 事件完成前，NLNP 的位置不在 NPL， 事件完成后，NLNP 的位置在 NPL。
III (“写”类)	+	+	+	(1) 动作所表达的事件可能在 NPL 发生， 也可能不在 NPL 发生； (2) 事件完成前，NLNP 或者不存在，或 者位置不在 NPL，事件完成后，NLNP 的位置在 NPL（换言之，事件的完成 对 NLNP 的位置发生显著影响）。
IV (“找”类)	+	-	-	(1) 动作所表达的事件可能在 NPL 发生， 也可能不在 NPL 发生； (2) 事件完成前，NLNP 的位置或者在 NPL， 或者不在 NPL，事件完成后，NLNP 的位置或者在 NPL，或者不在 NPL（换 言之，事件完成与否对 NLNP 的位置 没有显著影响）

需要说明的是：

(一) 一般情况下，跟“削”类和“找”类动词发生语义联系的“处所 NP”是“事件处所”角色，充任这个角色的 NP 必须出现在动词前位置，比如“张三在厨房里削土豆”，是指“削土豆”这个事件发生在“厨房里”，这个表达形式并不关心“土豆”所在的处所。但“削”类动词所表达的事件完成后，通常会造成实体的位置移动，比如“削”完之后，土豆的皮就从土豆表面剥离下来，这种“显著变化”，在句式表达上就通过格式 C 来实现。这时表示土豆皮所处处的“实体处所”角色必须放在动词后位置，比如“张三把土豆皮削在地上”。而“找”类动词所表达的事件发生后不会造成参与该事件的实体发生“显著的位置变化”，因而这类动词排斥“实体处所”角色与之共现。反映在句式上就是例 19 的 C 格式为不合法的表达形式。

(二) 跟“站”类和“写”类动词发生语义联系的“处所 NP”如果出现在动词前, 有时候可以理解为“事件处所”角色, 有时候可以理解为“实体处所”角色, 两种角色甚至可以重合, 因而造成相应的歧义句式; 如果出现在动词后, 则只能理解为“实体处所”角色, 不能理解为“事件处所”角色。不过, 出现在动词后位置的“实体处所”角色, 有时候仍然可以进一步分化为更细的语义角色, 比如“衣服放抽屉里”, 处所 NP“抽屉里”出现在动词“放”后, 应被看作为“实体处所”角色, 但这个句子可以分别变换为“衣服放在抽屉里”和“衣服放到抽屉里”两个不同句式¹², “抽屉里”在前一句中为“静态的实体处所”, 在后一句中为“动态的实体处所”, 或者可以理解为“表示终点的实体处所”角色(goal)。类似的例子还有“张三离开了北京”(其中的“北京”是“表示起点的实体处所”角色), “张三去了广州”(其中的“广州”是“表示终点的实体处所角色”), “张三站门口”, 同样也有歧义, “门口”可以理解为“静态的实体处所”角色(=“张三站在门口”), 也可以理解为“动态的实体处所”角色, 或“表示终点的实体处所”。值得指出的是, 例 16 中“站”类动词不能进入 C 格式, 并不是因为 NPL 成分不能放在动词后充任“实体处所”角色, 而是“把”字结构造成的, “把人站在门口”不能说, 是因为跟“把”字句表处置的句式语义要求相背, 如果把“把”拿掉, 就成为合法的表达形式。

(三) 从理论上讲, 以 A、B、C 三个格式为测试标准, 各取正负二值, 应该可以得到 $2^3=8$ 种不同情况, 也即分出八小类不同的动词。例 16—19 只列出了其中的四种情况。这里再补充两种情况: 一是 A、B、C 三个格式都排斥的动词, 比如“懂得”、“知道”等表示心理活动的动词(当然, 并不是所有的心理动词), 这类动词既不关心“事件处所”, 也不关心“实体处所”; 二是能进入 A、B 格式但不能进入 C 格式的动词, 比如“开会”、“唱戏”等, 这类动词只关心“事件处所”, 不关心“实体处所”。理论上还有两种情况, 就是 A、B 两格式中任意一个取负值, 剩下两格式同为正值的情形。但实际上没有这样的动词。限于篇幅, 对此就不展开讨论了。

以上句式变换示例的着眼点基本上都还是放在动词的“客体”论旨角色(一般居**动词后位置**的 NP)上, 下面的例子可以说明, 句式变换同样可以帮助判别**动词前位置**的 NP 所充任的不同的论旨角色。

(20) A	B	C
张三认识李四	—— * 张三把李四认识了	—— * 李四被张三认识了
张三支持李四	—— * 张三把李四支持了	—— * 李四被张三支持了
张三同意李四的观点	—— * 张三把李四的观点同意	—— * 李四的观点被张三同意
张三听说了很多谣言	—— * 张三把很多谣言听说了	—— * 很多谣言被张三听说了

(21) A	B	C
张三批评了李四	—— 张三把李四批评了(一顿)	—— 李四被张三批评了
张三吃了三个苹果	—— 张三把三个苹果吃了	—— 三个苹果被张三吃了
张三打了李四	—— 张三把李四打了	—— 李四被张三打了
张三请来了李四	—— 张三把李四请来了	—— 李四被张三请来了

例 20 和 21 中 A 组格式相同, 都是 X + V + Y 格式, 笼统地说, 其中出现在 X 位置上的 NP 都可以看作是“主体”, 出现在 Y 位置上的 NP 都可以看作是“客体”, 但如果把例 20A 和例 21A 分别变换为相应的 B 句式(“把”字句)和 C 句式(“被”字句), 就可以显现例 20 中的动词(“认识”类动词)和例 21 中的动词(“批评”类动词)实属不同的小类, 相应地, 出现在例 20 中的 X 位置上的 NP 和出现在例 21 中 X 位置上的 NP 也应视作充任了不同的论旨角色类型。比如, 可以把例 20 中的“张三”定为“感事”角色(不能出现

¹² 值得注意的是, 这种区分不同类型的“实体处所”的观念, 也有助于解释“他把信寄在邮局里”不合法, 而“他把信寄到广州”合法, 限于篇幅, 这里不详细讨论。

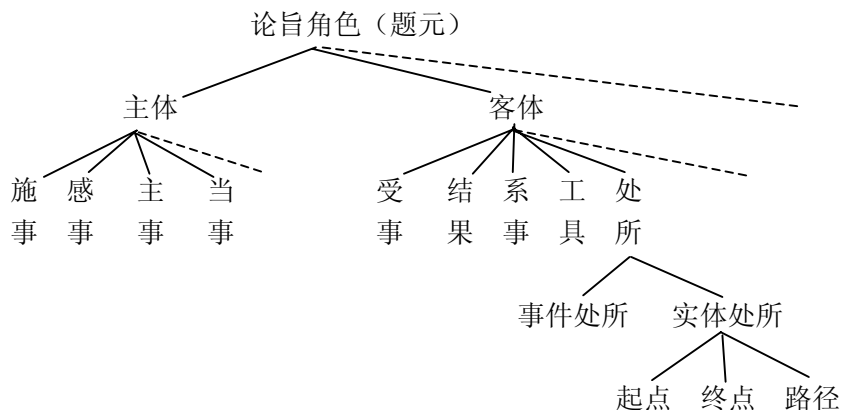
在“把”前位置，“被”后位置)；例 21 中的“张三”定为是“施事”角色(可以出现在“把”前位置，“被”后位置)。

动词前 NP 不是施事的例子还有如：

- (22) 汽车抛锚了 (可以回答“汽车现在怎么样了?”)
- (23) 电影院爆满 (可以回答“电影院现在怎么样了?”)
- (24) 数学及格了 (可以回答“数学现在怎么样了?”)

上面例子中动词前面的 NP 都不是“施事”“感事”，可以看作是“当事”¹³。当然，这些动词内部还可以根据句式变换性质的差异做进一步区分，比如“汽车抛锚了”可以变换为“抛锚的汽车”，“电影院爆满”可以变换为“爆满的电影院”，“数学及格了”却不能变换为“及格的数学”。可见它们相互之间还是有差异的。通过这些句式变换方面的系统的差异性来区分与动词同现的各个 NP 的论旨角色类型，同时也就对动词进行了更精细的小类划分。

本小节通过句式变换方式显示了一些论旨角色对不同句式的适应性差异。很显然，这种通过形式测试手段来区分语言范畴的不同类别的方式，跟我们使用分布测试手段来区分不同的词类(陆俭明，1994)，在操作模式上是完全一致的。如果能够按照这种方式对动词的论旨角色进行系统的研究，也必然会如词类系统一样，得到一个如下图所示的论旨角色的层级系统(作为示例，只是部分论旨角色，并不全面)。当然，因为语义范畴是相对的，这样的层级体系可以不只一个。不过，无论具体的范畴设置如何相对，但定义这些论旨角色的原则却都一样，那就是应该以句式变换上的形式特征为求同排异的准则。



四 句式变换的形式化规则

本节给出两个形式规则的实例，来说明论元和论旨角色等语义范畴在计算机处理自然语言中具体是如何起作用的¹⁴。

(25) VP → P<把> NP V P<在> NP

条件: %%NP.语义类=处所, %V.实体处所=%%NP, %V.受事=%NP OR %V.结果=%NP

转换: => p<在> np//NP v/V np/NP

分析式中的符号均用大写字母表示，转换式中的符号均用小写字母表示。“条件”部分的含义是，如果分析式中第二个 NP (用“%%NP”表示) 是“处所类”名词，动词的“实体处所”论旨角色的语义要求与第二个 NP 匹配 (用“%V.受事=%NP”表示)，“受事”或“结果”角色的语义要求跟第一个 NP (用“%NP”表示) 匹配，就执行相应的转换动作。

¹³ 注意这里的“电影院”也不是“处所”角色，它的前面加不上“在”这个格标，不能说“在电影院爆满”。

¹⁴ 这里所举规则及判断条件仅仅是展示以语义范畴为基础向计算机提供语言知识的一般方式而已，规则的准确性和覆盖能力并未经过一定规模语料的测试和检验。有关规则的详细讨论，可参看詹卫东 (2000)。

“转换”部分的含义是：将分析式中第二个 NP（用“/NP”表示）移动到动词前位置，将分析式中的第一个 NP（用“/NP”表示）移动到动词后位置。下面是实例：

- (26) 张三把房子盖在山下 → 张三在山下盖房子
张三把字写在纸上 → 张三在纸上写字
(27) 张三把书忘在宿舍里 → * 张三在宿舍里忘书
张三把日记本锁在抽屉里 → * 张三在抽屉里锁日记本

例 26 中动词“盖”“写”符合规则 25 中的条件，因而可以进行相应的转换，产生合法的表达；例 27 中动词“忘”“锁”都没有“实体处所”角色，不符合规则的条件要求，因而不能进行相应的转换。

(28) S → NP NP V U<了>

条件: %V.论元=1.5,%V.当事=%NP,%NP.语义类=有生命物

转换: => np/NP v/V u<了> np/NP

分析式中 U 代表助词，动词要求是“死”“断”等 1.5 价动词。同时要求第二个 NP 满足动词的“当事”论旨角色的语义要求，要求第一个 NP 是有生命类名词。满足条件的分析式经转换后，第二个 NP 移到动词后位置。下面是转换实例：

- (29) 张三父亲死了 → 张三死了父亲
张三眼睛瞎了 → 张三瞎了眼睛
(30) 张三媳妇生病了 → * 张三生病了媳妇
张三昨天瞎了 → * 张三瞎了昨天
张三弟弟抓到了 → * 张三抓到了弟弟

例 29 中动词“死”“瞎”符合规则对动词的论元要求，同时原式中两个 NP 也都满足动词的论旨角色要求和语义类要求，因此转换式是合法的表达形式。例 30 中“生病”“抓到”分别是 1 价和 2 价动词，不符合规则对动词的论元要求。“瞎”虽然符合论元要求，但“昨天”是时间名词，不符合动词对“当事”的论旨角色语义要求，因此这些原式都不满足规则的条件，因此无法进行转换。

把上述规则告诉给计算机，计算机在接收到例 26-27、例 29-30 的输入后，就能根据条件判断是否进行相应的转换。如果按照规则转换出合法的表达形式，并且转换式跟原分析式在逻辑真值上保持一致，人们就会“觉得”，计算机像是听懂了人的自然语言（至少是对上面例子中这很少的一部分自然语言，像是“懂”了）。

对于计算语言学的研究者来说，目标就是想办法帮助计算机只进行正确的转换，同时要排除错误的转换，而要达此目的，就要不断改进规则中的“条件”描述，通过设置更合理的句法语义范畴，使得条件约束更好地反映句式之间的变换关系，反映汉语的语言事实。

五 结语

传统面向人的语言学研究有一个基本的目标或者说是原则，即所谓“揭示形式与意义之间的对应关系”，那么，对计算机来说，这个目标或者原则也许可以换用一个说法，就是“揭示一个形式与另一个形式之间的对应关系”。只有把这样的对应关系系统地揭示出来，才有可能帮助计算机来模拟人的语言行为。因为只有按照形式特征归纳出的句法语义范畴，在描述规则的条件时才可以直接地派上用场，而“句式变换”可以帮助突显语言成分的“形式”特征，因此本文基于上述认识，以动词论元结构为考察对象，从句式变换上的差异来寻求形式上具有可操作性的依据去刻划动词的论元性质以及论旨角色性质，本文所做的初步探索工作说明，至少就面向计算机的语法研究而言，这样做是合理的。

事实上，语言学界一直以来就强调将“变换分析”作为一种基本的语法分析方法，并通

过这种手段取得了不少分化歧义句式的研究成果（朱德熙，1986；袁毓林，1989，陆俭明，1990），本文则试图进一步强调，“句式变换”本身就应该看做是一种广义的形式特征，换句话说，当我们打算刻画“施事”这样的语义范畴，区分不同的名词语义角色，以及在语义层面更细致划分动词小类时，不应该仅仅局限在一个格式内部观察，而应该把视野拓宽到多个格式之间的系统的变换关系上，在更大的范围中寻找形式特征，这样才能发掘出更多更准确的动词小类以及动名关系类型。

当然，在强调“形式标准”的同时，我们也不反对从“意义”上去认识动词的“论元性质”和“论旨角色”这些概念。跟“名词”等概念一样，我们也可以说“名词是表示事物的词语”，这是从语法意义上对“名词”所做的说明，只不过，计算机在把一个词标记为“名词”时，它的“语法意义”是次要的，而该词语具有一系列的名词的“形式特征”是主要的，正是这些“特征”决定了这个词语在实际使用时的“行为表现”，直接影响计算机的分析结果。同样的，我们把一个动词标记为 X 元动词时，也可以从意义上来说明该动词可以跟 X 个名词性成分发生关联，但更重要的是，这意味着该动词具有一系列的形式特征（比如各种句式变换特征），可以作为计算机判断相关的句式变换是否合法的形式依据。再比如，像“施事”这个概念，当然可以从“动作的发出者，实施者”这个角度去加以认识，但作为更具操作性的定义，则应该从形式特征的角度去描述，给一个 NP 贴上“施事”的标签，将直接制约该 NP 在一系列的句式变换中有不同于其他 NP 的“行为”特征。

参考文献

- 顾阳（2000）《论元结构及论元结构变化》，沈阳 编《配价理论与汉语语法研究》，语文出版社。
- 郭锐（1995）《述结式的配价结构与成分的整合》，沈阳等编《现代汉语配价语法研究》，北京大学出版社。
- 李临定（1988）《汉语比较变换语法》，中国社会科学出版社。
- 鲁川 等（2000）《交易类四价动词及汉语谓词配价的分类系统》，《汉语学习》2000 年第 6 期。
- 陆俭明（1990）《变换分析在汉语语法研究中的运用》，《湖北大学学报》1990 年第 3 期。
- 陆俭明（1994）《关于词的兼类问题》，《中国语文》1994 年第 1 期。
- 沈家煊（2000）《句式和配价》，《中国语文》2000 年第 4 期。
- 沈阳（1994）《现代汉语空语类研究》，山东教育出版社。
- 王占华（2000）《“吃食堂”的认知考察》，《语言教学与研究》2000 年第 2 期。
- 徐杰（2001）《普遍语法原则与汉语语法现象》，北京大学出版社。
- 徐烈炯（1995）《语义学》（修订版），语文出版社。
- 徐烈炯、沈阳（1998）《题元理论与汉语配价问题》，《当代语言学》1998 年第 3 期。
- 袁毓林（1989）《论变换分析法》，《汉语学习》1989 年第 1 期。
- 袁毓林（1998）《汉语动词的配价研究》，江西教育出版社。
- 詹卫东（2000）《面向中文信息处理的现代汉语短语结构规则研究》，清华大学出版社、广西科学技术出版社。
- 詹卫东（2001）《确定语义范畴的原则与语义范畴的相对性》，《世界汉语教学》2001 年第 2 期。
- 张伯江（1999）《现代汉语的双及物结构式》，《中国语文》1999 年第 3 期。
- 朱德熙（1978）《“在黑板上写字”及其相关句式》，《语言教学与研究》1978 年第 3 期。
- 朱德熙（1986）《变换分析中的平行性原则》，《中国语文》1986 年第 2 期。
- Goldberg, Adele E. (1995), *Construction: A Construction Grammar Approach to Argument Structure*, The University Chicago Press.
- Nerbonne, John (1996), *Computational Semantics -- Linguistics and Processing*, In Shalom Lappin, ed. *The Handbook of Contemporary Semantic Theory*, Oxford: Blackwell, 1996, Chapter 17.

Argument Structure And Constructional Transformation

Zhan Weidong

Department of Chinese Linguistics and Literature, Peking University

Abstract This paper surveys the argument structure of Chinese verbs from the viewpoint of constructional transformation. The author argues that the semantic categories, such as argument and theta role, should be defined depending on formal criteria of constructional transformations. In fact, "X-place" verb stands for sub-classes of verbs. "X" should be regarded as a symbol instead of a number. Even though it is regarded as a number, the value of X should NOT be limited in the range of 0,1,2,3, but can be expanded to more values, including 1.5, 2.5, etc.. "X-place" verb means both how many NPs can co-occur with the verb and where the NPs can be positioned. Furthermore, the theta roles, such as agent, location, can also be used to predict whether a constructional transformation is acceptable or not. If those theta roles can be classified and defined according to how they influence the constructional transformations systematically, the semantic categories will be very useful in helping the natural language processing by computer.

Keywords argument structure, theta role, verb, constructional transformation, formalized rule, constraint