

HPSG 理论简介 *

吴云芳 北京大学

提要 本文概要介绍了 HPSG 理论 (中心语驱动的短语结构文法)。表层导向、基于约束和词汇主义是 HPSG 的主要理论主张;特征结构是 HPSG 描述语法信息的主要手段;词汇类体系结构和词汇规则构成了 HPSG 的词汇体系;中心语——补足语规则、中心语——指定语规则等以及中心语特征原则、值传递原则等构成了 HPSG 的规则系统。文章最后对 HPSG 和 GB 理论进行了简要的比较。

关键词 HPSG 句法理论 计算语言学

0. 前言

HPSG (Head-Driven Phrase Structure Grammar), 即“中心语驱动的短语结构文法”, 是现今西方很为流行的一种理论。Pollard and Sag (1987) 首先勾勒出了 HPSG 的原型, 然后在 1994 年对其进行了一些改进, 从而形成了一个完整的句法理论。今天, 依据 HPSG 进行理论研究和依据 HPSG 进行计算语言学实践的都颇有学者在。后者如在德语、日语、英语三国语之间进行实时翻译的庞大的 Verbmobil 工程。HPSG 研究的大本营在 Stanford 的 CSLI (Center for the Study of Language and Information)¹。从 1994 年起, 每年都有关于 HPSG 的国际会议召开, 2003 年将是第十届会议²。

象其他句法理论一样, HPSG 也是在不断的变化着、发展着。本文的叙述基本上依据 Sag and Wasow (1999) 的版本。

1. HPSG 的理论主张

HPSG 是基于约束的词汇主义, 它来源于这样的心理语言学事实: 语言理解是以一种

* 本文研究工作得到国家 973 项目 (G1998030507-4, G1998030504-01) 和 863 项目 (2002AA117010-08) 的支持。

¹ 本文的写作从 Stanford 的网站上查阅了很多资料, 网址是: <http://hpsg.stanford.edu/hpsg>。

² 历届会议的时间和地点分别为: 第九届, 2002, Seoul; 第八届, 2001, Trondheim; 第七届, 2000, Berkeley; 第六届, 1999, Edinburgh; 第五届, 1998, Saarbrücken; 第四届, 1997, Cornell University; 第三届, 1996, Pittsburgh; 第二届, 1995, Tubingen; 第一届, 1994, Copenhagen。第十届将于 2003 年在 Michigan 召开。

高度整合和渐进的方式进行的 (in a highly integrative and incremental fashion) , 语言的和非语言的因素掺和在一起共同参与语言理解的过程。具体讲, 基于约束的词汇主义包括下面三方面的内容:

1.1 表层导向 (Surface Oriented)

HPSG 关注的是具体的、表层的结构。它认为, 由词串的表层组成就可以直接推衍出一个句子是否合格。因此, 任何抽象的、深层的结构 (象空语类等) 在 HPSG 中是不存在的。表层导向的结果是 HPSG 只有单一的句法结构 (single level of structure), 而不像生成语法那样存在多层句法结构。

1.2 基于约束 (Constraint – Based)

HPSG 认为, 适合人类语言处理的语法应该是一个约束描述体系。HPSG 整个的构架是建立在约束机制上的, 词项 (lexical entry) 描述、语法规则、语法原则等都是通过约束来实现的。

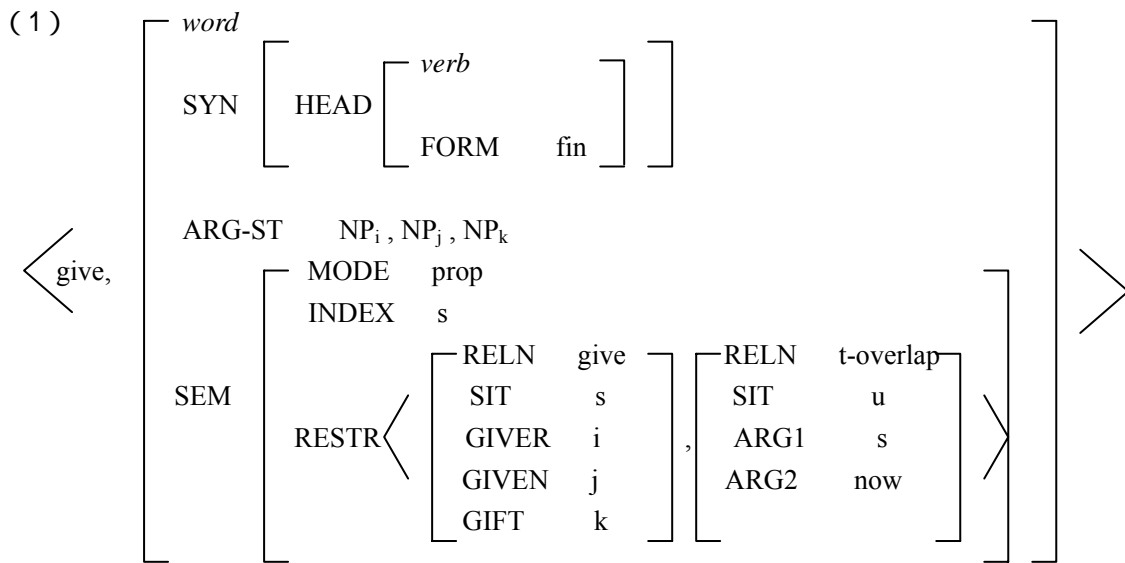
1.3 词汇主义 (Strict Lexicalism)

词语携带了丰富的句法语义信息, 它在很大程度上决定了它所在句子的句法语义结构。反过来, 句子之所以表现出不同的句法语义结构, 也正是因为其中所包含的关键词语不同。HPSG 把语法规则的重担几乎全部转移到了词汇上, 是严格的词汇主义。

2. HPSG 的特征结构

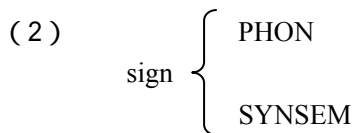
特征结构 (Feature Structure) 是描述语法信息的一种手段, HPSG 广泛采用复杂特征结构来描述词语或短语信息。复杂特征结构是 HPSG 实现其 “ 词汇主义 ” 的主要手段, 是 HPSG 理论的重要组成部分。

请看下面关于 “ give ” 的描写:



这里列出的是“give”主要的属性特征描述，并不是其全部。一个音义结合的语言符号 (sign)

最基本的属性特征是：



HPSG 不太关心 PHON (声音)的特征。SYNSEM 是 SYNTAX 和 SEMANTICS 的合写，

是 HPSG 主要关照的属性特征。(1) 描述的就是“give”的 SYNSEM 属性。

2.1 关于符号

[.....] 表示属性特征矩阵 (attribute-value matrix, AVM)，它是用来描述复杂特征结构的主要工具。

..... 表示属性特征的列表。 表示空表。

属性特征的特征值有两种形式：原子式 (atom) 和特征结构。上例中，“FORM fin”

是一个原子式的特征值，HEAD $\left[\begin{array}{l} \text{verb} \\ \text{FORM} \quad \text{fin} \end{array} \right]$ 是以特征结构为值的复杂特征结构。传统

的短语结构文法 (Phrase Structure Grammar, PSG) 多采用原子式的特征值，其描写能力受到

很大的限制。HPSG 大量采用复杂特征结构来表示特征值 (feature structure within feature

structure)，其描写能力大大加强了。

在下文的描述中，经常会看到属性特征值用 $1, 2, \dots, n$ 等表示，这称为“标签 (tag)”。相同的标签表示结构共享 (structure sharing)。结构共享的成分，是严格的字标一致 (token identity)，也就是完全合一。正是在这个意义上，HPSG 属于合一机制的文法。结构共享是 HPSG 核心的描述机制，正如“move”是管辖与约束理论 (GB) 核心的描述机制一样。

2.2 关于 SYN

SYN 用来描述语言符号的句法信息。HEAD 是其中最重要的一个属性，主要描述符号的词类 (part of speech)。FORM 表示符号的词形特征，此处“fin”表示该动词是“定式动词 (finite)”。

2.3 关于 SEM

SEM 用来描述语言符号的语义信息。HPSG 把句法和语义有机地结合在一起，在对符号的描述中包含了丰富的语义信息。HPSG 主要借鉴了情景语义学 (situation semantics) 的研究成果，其语义描写的目的是想说明：在一个事件中，谁对谁做了什么 (who did what to whom)，发生在什么时间、什么地点等。对一个符号的语义描写主要包括三个部分：

$$(3) \quad \begin{bmatrix} \text{MODE} \\ \text{INDEX} \\ \text{RESTR} \end{bmatrix}$$

MODE 有五个备选的属性值：prop (陈述)、ques (疑问)、dir (祈使)、ref (指称) 和 none。

例如，动词的 MODE 值是 prop，名词的 MODE 值是 ref。冠词、介词、连词等不能表达

陈述、疑问、祈使和指称的概念，其 MODE 值是 none。INDEX 对应于所描述的情景

(situation) 或事件 (event)。例如，在 (1) 的描述中，“give”传达了一个代号为 s 的事件。

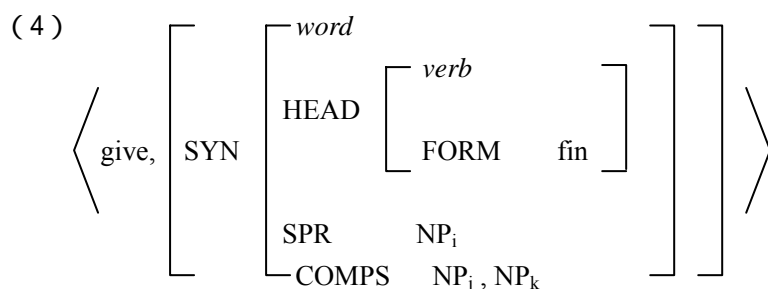
RESTR 指的是事件成立必须满足的一些条件 (conditions)。例如，“give”这一事件成立必

须满足这样的条件：情景 s 中 i 把 k 给了 j，发生的时间是现在，i、j、k 分别和 give 的

不同语义角色 (roles) 相联系。

2.4 关于 ARG-ST

ARG-ST (ARGUMENT-STRUCTURE) 用来描述语言符号的论元结构。它是一个属性特征列表,对动词而言,包含与动词共现的所有必需论元 (arguments), 是该动词的指定语 (SPR) 和补足语 (COMPS) 之和;对名词而言,主要是该名词的指定语 (SPR)。例如, (1) 中 give 的 ARG-ST 是 NP_i, NP_j, NP_k , 意思是它需要一个主语和两个补足语 (一个间接宾语和一个直接宾语) 与其共现。在 ARG-ST 中, 论元的排列次序和实际句子中的次序是相吻合的: “ Mary(NP_i) gives me(NP_j) a book(NP_k) ”。 NP_i 是 $NP [SEM [INDEX \ i]]$ 的省写, NP_j 、 NP_k 类似。 NP_i 、 NP_j 、 NP_k 分别对应于 RESTR 中的 GIVER、GIVEN 和 GIFT, 这样, give 的句法结构和语义特征就有机地联系起来。有时, 可以不用论元结构 ARG-ST 而分别用指定语 SPR 和补足语 COMPS 来描述, 例如对 “ give ”:



2.5 特征结构是一个严密组织的体系

特征结构是经过分类排序的 (be sorted)。首先, 每一个类都有自己独特的特征值, 不同的类要求不同的特征值。例如, SEM、SYN 是不同的类 (sort)³, SEM 要求 MODE、INDEX、RESTR 作为它的值, 而 SYN 要求 HEAD 作为它的值的一部分。再者, 类是一个有层级的体系结构。例如, PHON 和 SYNSEM 是 sign 的下层类 (subsort), HEAD 是 SYN 的下层类。

2.6 一点说明

对一个符号的特征结构描述并不一定要求是完全的, 可以根据需要择其相关, 事实上,

³ “ sort ” 的概念和下文的 “ type ” 是一样的, 都相当于汉语中的 “ 类 ”。“ typed feature structure ” 和 “ sorted feature structure ” 也基本上是一样的, 都是 “ 类特征结构 ”。

本文的特征结构描述都是不完全的。HPSG 的特征结构异常复杂，本节介绍的只是其中主要的部分，下面我们将根据需要随文再介绍一些。

3. HPSG 的词汇

HPSG 把语法描写的重担转移到了词汇上，一个词语将不得不承担太多的句法和语义信息，这样词汇系统将不可避免地变得冗余和繁琐。HPSG 利用两个策略成功地解决了这一难题：词汇类体系结构（type hierarchy）和词汇规则（lexical rules）。

3.1 词汇类体系结构

HPSG 采用了类特征结构（Typed Feature Structure）。语言中的不同实体——语音、词、短语、句子等都是不同的“类（type）”，分别要求不同的属性特征与其相对应。语言中客观存在一个词汇类体系结构。

3.1.1 上层类（supertype）和下层类（subtype）

如果 T_2 是 T_1 的下层类，那么：

- a、适合于 T_1 的每一个特征同样适合于 T_2 ；
- b、与 T_1 相关的每一个约束⁴都影响到 T_2 的实例（instances）和下层类。

a、b 所规定的约束承袭（the inheritance of constraints）是单调的（monotonic），即承袭过程中不允许有例外发生，这与语言事实是不符的。例如，一般的英语名词不需要有“CASE”的属性，但作为其下层类的代词却要求“CASE”的属性；一般的英语动词只有第三人称和非第三人称两种形态变化，但作为其下层类的“be”却有第一人称、第二人称、第三人称三种形态变化。

3.1.2 约束的缺省承袭（default inheritance of constraints）

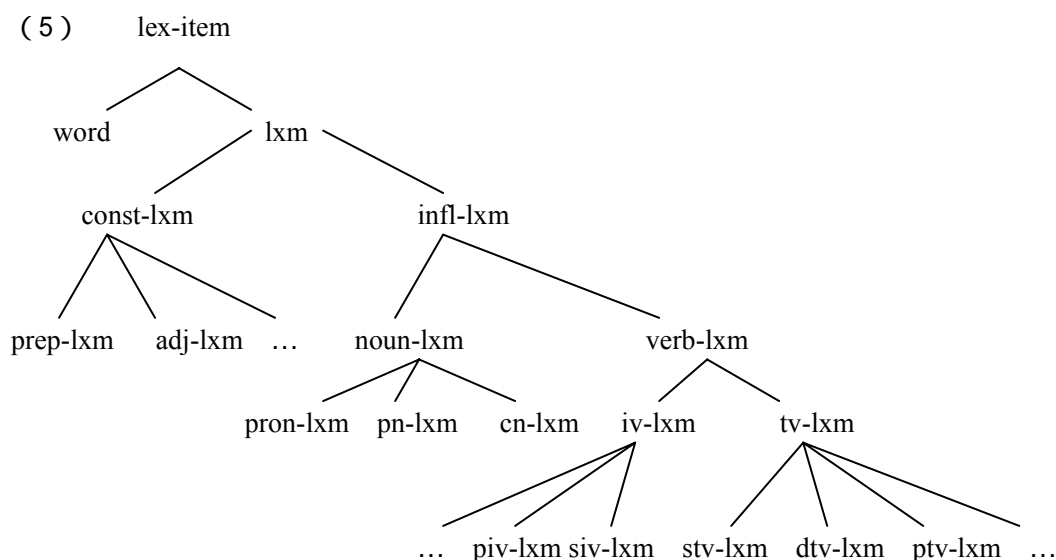
语言有规则，但规则有例外。约束规则有两种：一种是不可违背的（inviolable），下层

⁴ 此处的约束不是我们平常理解的狭义约束，它包含很多信息：特征的值，以及特征值之间的联系等。

类自动地、无例外地从上层类承袭；一种是缺省的。约束的缺省承袭是指：上层类缺省的约束规则可以被下层类特殊的、例外的约束规则覆盖和否定 (overridden)。

3.1.3 示例：一个英语的词汇类体系结构

Sag and Wasow (1999) 建立了一个英语的词汇类体系结构, (5) 是其中的部分：



下面对其作些说明。

lex-item (lexical-item) 是最上层的一个类。

word 和 lxm (lexeme) 是 lex-item 的直接下层类。word 指有词形变化的音义结合体, lxm 是不关注词形变化的一个词语“家族”, 是一个抽象的、静态的原型词(proto-word)。例如, walk, walks, walked 是不同的 word, 但属于同一个 lxm。lxm 是语言描写的出发点, word 是由 lxm 演化而来的。

根据是否有形态变化, lxm 分为 infl-lxm (inflecting-lexeme) 和 const-lxm (constant-lexeme)。

const-lxm 进一步分成 prep-lxm, adj-lxm 等。

infl-lxm 分成 noun-lxm 和 verb-lxm。

noun-lxm 进一步分成 pn-lxm (proper-noun-lexeme, 专有名词), cn-lxm

(common-noun-lexeme, 普通名词) 和 pron-lxm (pronoun-lexeme, 代词)。

verb-lxm 分成 iv-lxm (intransitive-verb-lexeme) 和 tv-lxm (transitive-verb-lexeme)。

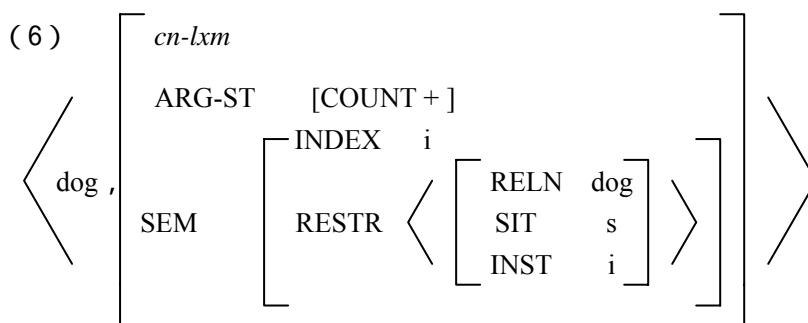
iv-lxm 进一步分成 siv-lxm (strict-intransitive-verb-lexeme) 和 piv-lxm (prepositional-intransitive-verb-lexeme)。前者指不带任何宾语的不及物动词, 如 “run”; 后者指带介词结构宾语的不及物动词, 如 “rely”, “I rely on you (我依靠你)”。

tv-lxm 进一步分成 stv-lxm (strict-transitive-verb-lexeme), 如 “devour”; dtv-lxm (ditransitive-verb-lexeme), 如 “give”; ptv-lxm (prepositional-transitive-verb-lexeme), 如 “put”。

3.1.4 词汇类体系结构的运作

词汇类体系结构从两个方面简化了 HPSG 的词汇操作:(1) 一定的类对应一定的属性特征, 因此没必要去描述名词的及物性或形容词的人称特性;(2) 类及其属性特征是一个有规律的层级体系, 只要知道了一个符号在整个体系结构中的位置, 它的大部分的句法和语义信息就可以自动地获得, 而没必要繁琐地逐一去描写。

在词汇类体系结构中, 一个完整的词位描述 (lexeme description) 包括两部分: 一是词位基础信息, 一是从上层类承袭来的信息。例如, 要对 “dog” 进行描述, 首先是它的基础信息:



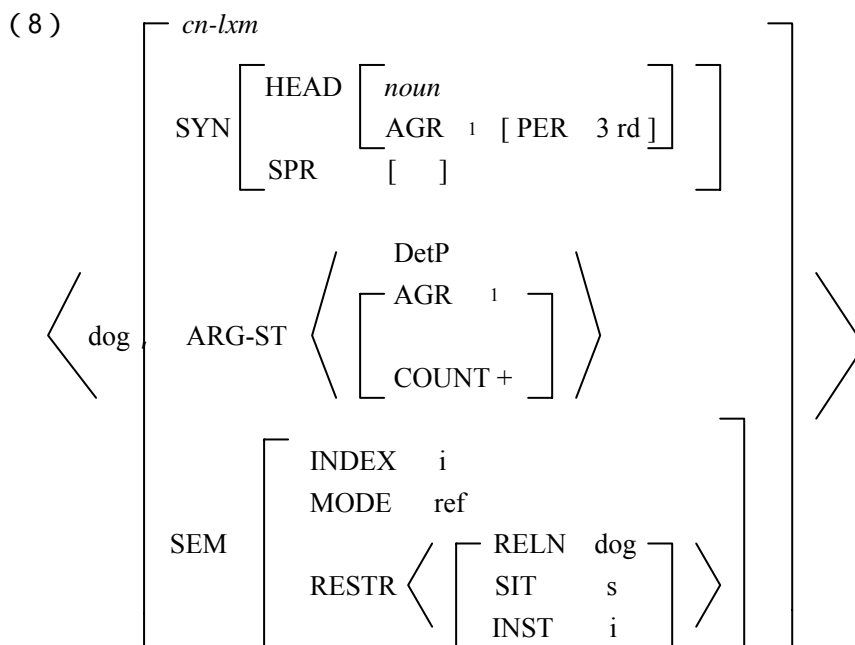
COUNT 是为了描述英语中的一致关系 (agreement) 而引入的一个属性特征, ARG-ST

[COUNT +] 表示 dog 要求一个可数的指定语 (如 “the, a” 等) 与其共现。dog 属

于 cn-lxm (普通名词) 类, 它要从上层类承袭一些信息来为己所用。其承袭轨迹如下:

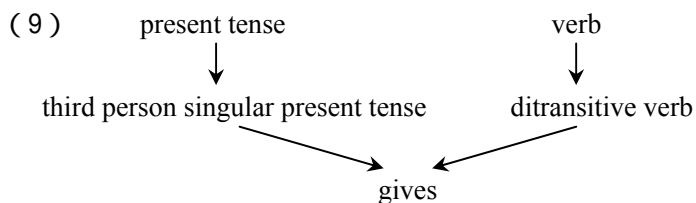
(7) lex-item → lxm → infl-lxm → noun-lxm → cn-lxm — dog

承袭过程中要遵循上层类、下层类的约定和约束的缺省承袭机制。把 dog 自己的基础信息和承袭来的信息相加，我们就得到了 dog 的属性特征结构：



3.1.5 多亲承袭体系 (multiple inheritance hierarchy)

单亲承袭体系 (single-mother inheritance) 是一个儿子节点只从一个母亲节点处获取信息，它不能有效地描述跨类的语言信息。对一个符号，可以从不同的角度对其进行分类，不同角度的分类对应不同的信息维度。HPSG 更经常用到的是多亲承袭体系，即一个儿子节点从多个母亲节点处获取信息。多亲承袭体系无疑比单亲承袭体系具有更大的概括性和更强的描写能力。用一个简单的例子来说明：



“gives”从“现在时”处获取“第三人称单数现在时”的信息，从“动词”处获取“双宾语及物动词”的信息，当然还可以从别处获取相关信息，所有这些信息相加共同成为“gives”

的词汇信息。

3.2 词汇规则

词汇规则是一个产生式的装置，它具有这样的形式： $X \longrightarrow Y$ 。X 是输入，经过规则运算，输出 Y。输入的 X 是一个完整的词位描述（complete lexeme description），即经过词汇类体系结构运作后、包含有完整信息的词位描述。

HPSG 有两类不同的词汇规则，下面分别说明。

3.2.1 形态规则（Inflectional Rules）

这是有关形态变化的规则，说明如何从一个词位产生具有形态变化的词项（from lexeme to lexical entry）。HPSG 有很多形态规则：单数名词规则、复数名词规则、动词过去式规则、动词进行式规则等等。这儿举复数名词规则为例来说明。

$$(10) \left\langle 1, \left[\begin{array}{l} \textit{noun-lxm} \\ \text{ARG-ST} \quad [\text{COUNT +}] \end{array} \right] \right\rangle \longrightarrow$$
$$\left\langle F_{NPL}(1), \left[\begin{array}{l} \textit{word} \\ \text{SYN} \quad \left[\text{HEAD} \quad \left[\text{AGR} \quad [\text{NUM} \quad \textit{pl}] \right] \right] \right] \right] \right\rangle$$

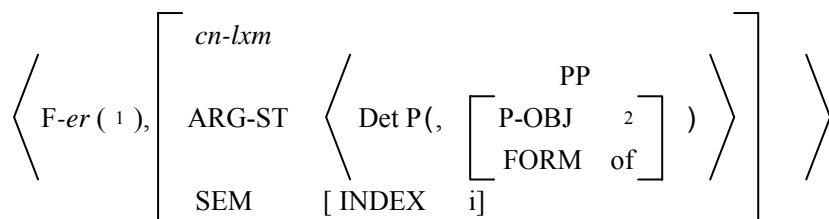
规则的输入是一个 *noun-lxm*，其指定语的特征是 [COUNT +]，这样就排除了像“furniture”这样的不可数名词。F_{NPL} 是一个形态函数，将一个名词由原形变成复数形式。输出的类由输入的 *noun-lxm* 变成了 *word*，并且具有了属性 [NUM pl]。

3.2.2 派生规则（Derivational Rules）

派生规则说明如何从一个词位产生与其相关的另一个词位（from lexeme to lexeme）。看

下面这个名物化名词的词汇规则：

$$(11) \left\langle 1, \left[\begin{array}{l} \textit{verb-lxm} \\ \text{ARG-ST} \quad \text{NP}_i (, \textit{ }^2 \text{ NP}) \\ \text{SEM} \quad [\text{INDEX} \quad \textit{s}] \end{array} \right] \right\rangle \longrightarrow$$



ARG-ST 中的“()”表示可选,“ARG-ST NP_i (, 2 NP)”表示输入是一个不及物动词或及物动词,但不能是一个带介词结构做宾语的动词,因此不能有 *rely* → *relyer*。

前后的“类”发生了变化,由输入的 *verb-lxm* 变成了输出的 *cn-lxm*。F-er 是一个形态函数,在动词词尾附加“er”。输入动词的指定语和输出名词具有相同的 INDEX (都是 i),这是容易理解的:在一个驾驶的情景 s 中,drive 的指定语和 driver 是同一的。下面的例子可以帮助理解输出中的 PP:

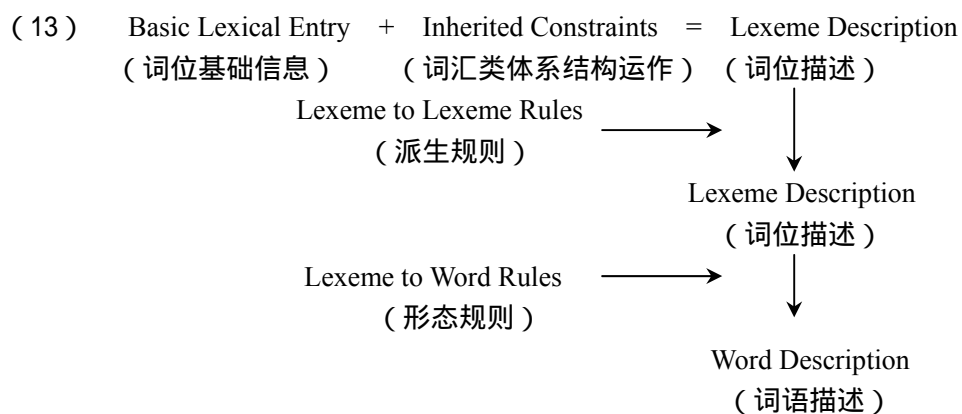
(12) He discovered the oxygen → the discoverer of oxygen

“oxygen”由“discover”的补足语变成了介词“of”的宾语,即(11)中标签同为 2 的那个成分。

词汇规则简化了 HPSG 的词汇操作,使词的描述更为简捷。另一方面,利用词汇规则,HPSG 可以有效地分析和解释一些句法现象,而无需借助转换,实现了它的“词汇主义”。

3.3 词汇体系的运作

有了词汇类体系结构和词汇规则,HPSG 的词汇操作大大得到了简化。HPSG 词汇体系的运作可直观地描述如下:



4. HPSG 的规则

HPSG 的规则都是围绕中心语展开的，正是在这个意义上它被叫做“中心语驱动”（head-driven）。HPSG 认为，传统的 PSG 没有认识到中心语的存在，因而缺乏概括性（generalization）。缺乏中心语，是 HPSG 挑战传统 PSG 的主要方面之一。HPSG 中，短语和词语的描写在形式上基本是一样的。不同的是，短语有“子女（daughter）”的属性，其中有一个是中心子女（head-daughter），其余的子女可依据功能分别叫做“补足语子女（complement-daughter）”、“修饰语子女（modifier-daughter）”等。中心语规则都具有这样的形式：

(14) [phrase] → H[].....

象其他的句法理论一样，HPSG 的规则系统也有原则（principle）和规则（rule）两部分组成。

4.1 规则

Pollard and Sag (1994) 用直接支配模式（Immediate Dominance Schemata IDS）来描述规则，分为模式一（Schema 1）、模式二、模式三等。Sag and Wasow (1999) 分为不同的规则来进行描述。两者其实是统一的，而后者显得更为直观和简洁，因此本文采用后者的方法。

4.1.1 中心语 — 补足语规则（Head - Complement Rule）

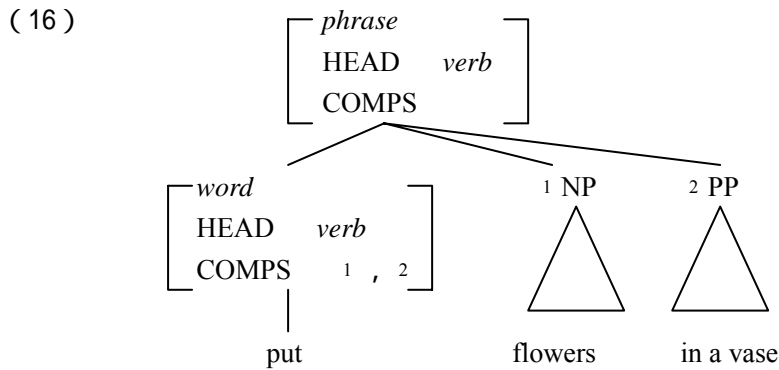
补足语指中心语句法上要求同现的成分⁵，HPSG 用 COMPS 的属性来表示。COMPS 是一个属性特征列表，其成分的排列次序和实际句子中的次序相吻合。例如，put 的 COMPS 值是 NP, PP⁶，give 的 COMPS 值是 NP, NP。COMPS 中用“()”表示可选，如 eat [COMPS (NP)]；用“|”表示或取，如 deny [COMPS NP|S]。规则如下：

⁵ 与中心语同现的成分是由句法决定的还是由语义决定的，并没有定论。HPSG 把它看成是一种句法现象。

⁶ PP 什么时候是 COMPS 的一部分，什么时候不是，HPSG 中有学者专门对其进行过研究。

$$(15) \left[\begin{array}{l} \textit{phrase} \\ \text{COMPS} \end{array} \right] \rightarrow H \left[\begin{array}{l} \textit{word} \\ \text{COMPS} \quad 1 \dots n \end{array} \right] \quad 1 \dots n$$

中心语 — 补足语规则要求所有的补足语实现为中心语的兄弟节点。如 put :



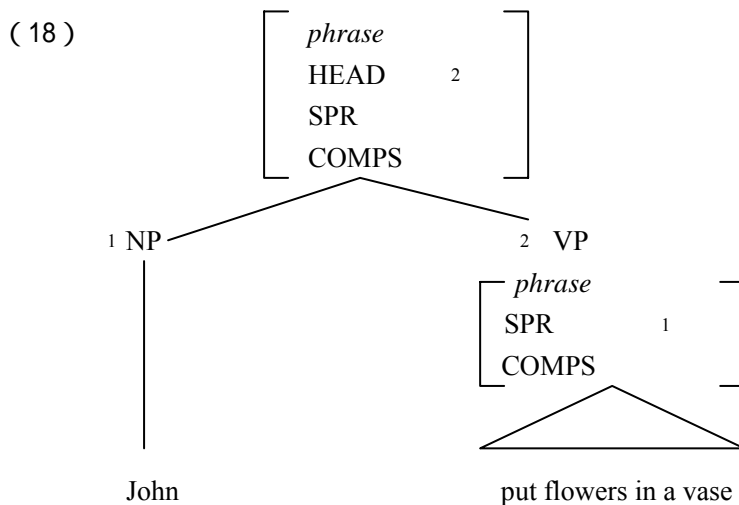
4.1.2 中心语 — 指定语规则 (Head — Specifier Rule)

指定语指动词中心语的主语 (subject), 名词中心语的限定性成分 (determiner), HPSG

用 SPR 的属性来表示。同 COMPS 一样, SPR 也是一个属性特征列表。规则如下:

$$(17) \left[\begin{array}{l} \textit{phrase} \\ \text{COMPS} \\ \text{SPR} \end{array} \right] \rightarrow 1 \quad H \left[\begin{array}{l} \textit{phrase} \\ \text{SPR} \quad 1 \end{array} \right]$$

需要注意的是,在中心语 — 指定语规则中,中心语子女的类型是 *phrase*;而在中心语 — 补足语规则中,中心语子女的类型是 *word*。中心语 — 补足语规则发生在中心语 — 指定语规则之前,即中心语总是先和补足语捆绑,然后作为一个 *phrase* 再和指定语捆绑。如对“John put flowers in a vase”的分析:



4.1.3 中心语—修饰语规则 (Head — Modifier Rule)

修饰语指修饰中心语的成分，HPSG 用 MOD 的属性来表示修饰语的修饰功能，MOD 同样是一个属性特征列表。例如形容词的 MOD 属性是 [MOD NP]，副词是 [MOD VP]。

其规则是：

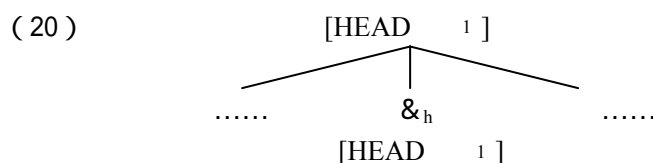
$$(19) \quad [phrase] \longrightarrow H \ 1 \ [phrase] \quad \left[\begin{array}{l} phrase \\ MOD \ 1 \end{array} \right]$$

除了上面三条主要的规则外，HPSG 还有并列结构规则（Coordination Rule）等，本文不加详述。

4.2 原则⁷

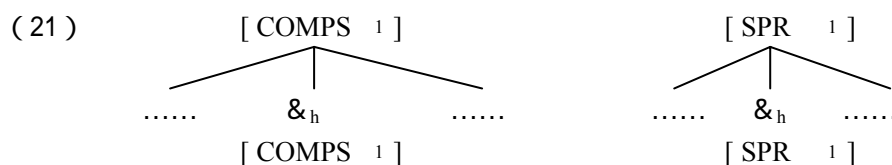
4.2.1 中心语特征原则（Head Feature Principle HFP）

HPSG 中，最基本的、最普遍通用的原则是中心语特征原则：任何一个中心语短语中（headed phrase），母亲的中心语特征值和中心子女的中心语特征值合一。其形式描述是：



4.2.2 值传递原则（Valence Principle）

为了使母亲的指定语 SPR 和补足语 COMPS 属性和中心子女保持一致，HPSG 引入了值传递原则：除非特别说明，母亲的 SPR 和 COMPS 的特征值和中心子女同。图示如：



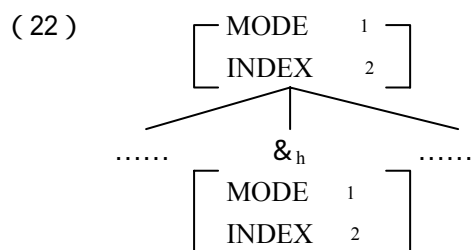
以上两个原则是关于句法的，下面两个原则是关于语义的。

4.2.3 语义承袭原则（Semantic Inheritance Principle）

这个语义原则同样体现了“中心语驱动”：在一个中心语短语中，母亲的 MODE 和

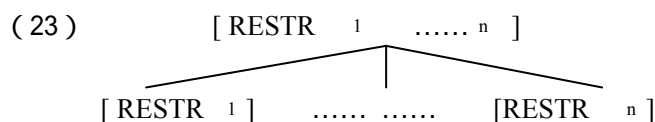
⁷ 之所以把原则放在规则之后叙述，是因为原则中要用到 COMPS、SPR 等概念。

INDEX 特征值和中心子女同。图式如：



4.2.4 语义组合原则 (Semantic Compositionality Principle)

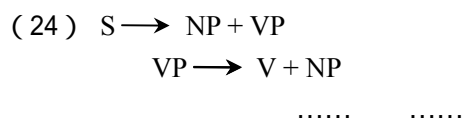
在一个短语结构 (well-formed phrase structure) 中，母亲的 RESTR 值等于所有子女的 RESTR 值之和。图式如 (HPSG 中用符号 \oplus 表示“和”)：



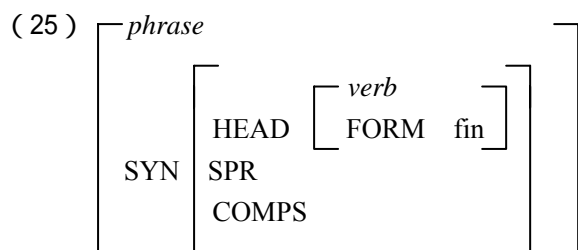
除了上面四个原则外，HPSG 还有论元实现原则 (Argument Realization Principle)、回指一致性原则 (Anaphoric Agreement Principle)、约束原则 (Binding Principle) 等。本文不加详述。

4.3 起始符 (Initial Symbol)

上下文无关文法 (Context -Free Grammar, CFG) 的短语结构规则是从句子 S 开始的，下面是我们所熟悉的规则：



HPSG 理论中没有 S 的概念，它的起始符是一个满足了一定约束条件的 *phrase*：



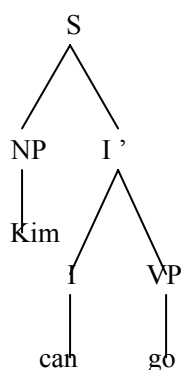
首先，这个 *phrase* 的中心语是一个 *verb*，它的形式是有定的 (finite)。其次，它是完全饱

和的 (saturated), SPR 和 COMPS 的值都为空。这样来定义起始符, HPSG 的整个理论框架变得更加完美和统一。

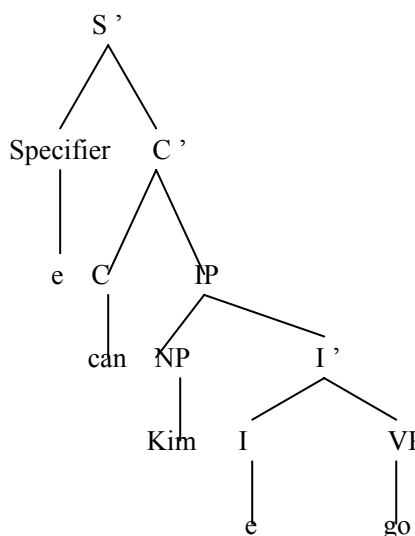
5. HPSG 和生成语法

Sag and Wasow (1999 : 411) 指出, HPSG 是从 Chomsky 生成语法模式中发展而来的一种理论, 两者有很多相似之处, 例如, HPSG 的中心语特征原则就酷似 GB 的投射原理。但总起看来, HPSG 和生成语法 (以 GB 为代表) 是很不相同的两种句法理论。它们之间最主要的差别是: HPSG 是非转换的 (nonderivational), 表现出来的是一个平面的单树结构 (flat structure); 而生成语法存在大量的移位、转换现象, 表现出来的是一个多层面的多树结构。正是在这一点上, HPSG 比生成语法更为计算语言学界所青睐。我们举个简单的例子来说明。譬如, 要对助动词移位进行解释: Can Kim go? GB 是这样刻画的:

(26) D-structure :

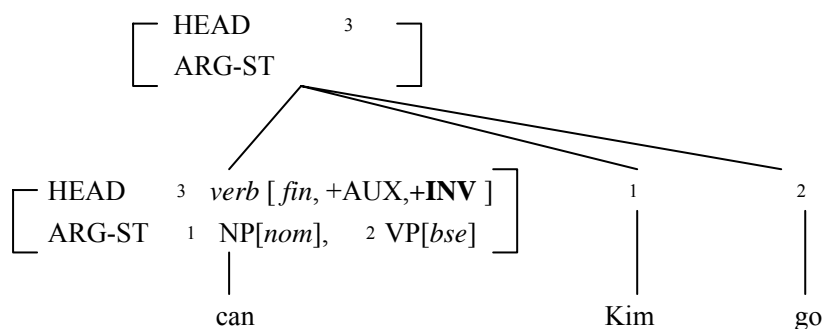


S-structure :



HPSG 是这样刻画的 :

(27)



HPSG 在动词中心语的 HEAD 属性中引入了 INV(INVERTED)的特征。在符合 HPSG 原则和规则的前提下,具有 [+INV] 特征的动词可以移位。HPSG 用词语的一个属性特征实现了 GB 的移位。

6. 结束语

HPSG (中心语驱动的短语结构文法)是一个复杂的、不断前进着的句法理论。本文介绍的只是 HPSG 基本的、核心的内容,更多具体的东西我们并没有涉及。HPSG 对一致性问题 (agreement)、补足语结构 (complement structure)、约束 (binding) 问题等都提出了自己的解决方案。HPSG 最新的发展,是建立“基于符号的概念语法 (sign-based conception of grammar)”,整个体系变得愈发抽象,也更具概括力。

像生成语法一样,HPSG 同样旨力于普遍语法 (universal grammar) 的探讨。那么,HPSG 在多大程度上适合于汉语呢?汉语和英语等印欧语言存在着很大的差别,它是一种没有形态变化的语言,HPSG 中的词汇形态规则、对一致性问题 (主语-谓语一致,指定语-名词中心语一致) 的解决方案似乎在汉语中无用武之地。HPSG 中围绕中心语的一系列原则和规则是否可以在汉语中畅通无阻地加以应用呢?

很多学者已经感受到了“词汇主义”的思潮,纷纷投入精力和财力去进行词汇研究。我们是否可以借鉴 HPSG 的词汇类体系结构,来简化繁复的词汇描述呢?

参考文献

- Anthony, D. 2001. *Linking by Types in the Hierarchical Lexicon*. Chicago: The University of Chicago Press.
- Borsley, R. D. 1996. *Modern Phrase Structure Grammar*. Cambridge: Blackwell Publishers.
- Chomsky, N. 1981. *Lectures on Government and Binding*. Dordrecht :Foris.
- Nightingale, S. 1998. Book Review. *Machine Translation* 13: 229-232.
- Pollard, C. and I. A. Sag. 1987. *Information Based Syntax and Semantics* . Chicago: The University of Chicago Press.
- Pollard, C. and I. A. Sag. 1994. *Head-Driven Phrase Structure Grammar* . Chicago: The University of Chicago Press.

Sag, I. A. and T. Wasow. 1999. *Syntactic Theory: A Formal Introduction*. Center for the Study of Language and Information, Stanford, California.

作者通讯地址：100871 北京大学计算语言学研究所

Abstract This paper gives a brief introduction to HPSG. The leading ideas of HPSG are surface oriented, constraint-based and strict lexicalism. Feature structure is the main method to describe syntactic information. Type hierarchy and lexical rules constitute HPSG's lexical system. HPSG's rules include *head-complement rule*, *head-specifier rule* and so on, and it has principles such as *head feature principle*, *valance principle* and so on. This paper also makes a brief comparison between HPSG and GB in the last section.

Key Words HPSG syntactic theory computational linguistics