

句法学导论

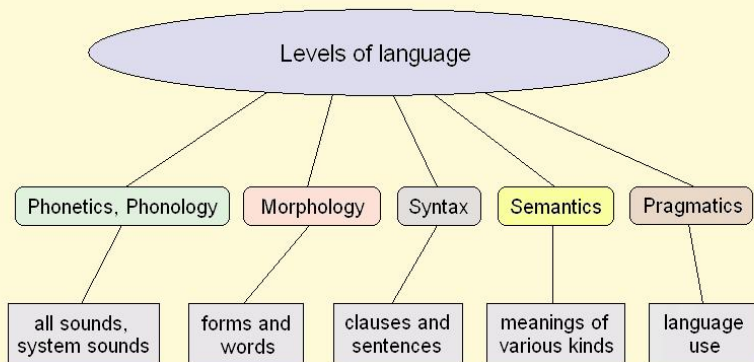
Anqi Zhang

Nanjing University

Introduction to Linguistics

10/21/2022

句法



句法范畴

句法研究词与词如何连接成句

词语在句子中出现的位置很大程度由其句法范畴（**syntactic category**）决定，句法范畴又称词类（**part of speech**）

语法关系（**grammatical relation**）描述短语之间如何互相联系

词类 vs. 语法关系

Parts of Speech	Grammatical Relation
noun	subject
verb	object
adjective	predicate
adverb	modifier
...	...

词汇范畴

形态句法范畴（Morphosyntactic）可以分为词汇范畴（**lexical**）和**功能**范畴（**functional**）：

词汇范畴：

Noun *dog, party, book, Fred, happiness*

Verb *go, run, see, give, rain, blacken*

Adjective *narrow, black, happy, bookish, rainy*

Adverb *happily, narrowly, well*

功能范畴

介词 (Preposition) *in, around, of, about, between*

限定词 (Determiner) *the, a, some, any, that, this*

标补语 (Complementizer) *that, if, whether*

并列连词 (Coordinator) *and, but, or, nor*

程度修饰词 (Degree modifier) *very, really, too*
that, more, less, as

助动词 (Auxiliary) *would, could, might, have, be*

代词 (Pronoun) *we, him, I, they, us, her, its*

词汇范畴与功能范畴

词汇范畴基本上都是**开放范畴**（**open categories**），也就是说，很容易构造出属于这些范畴的新词（例如*iPod*, *google*, or *defriend*）

功能范畴则是**封闭范畴**（**closed categories**），也就是说，很难构造出属于这些范畴的新词

词汇范畴与功能范畴

词汇性词语的意义往往很具体，容易把握；这些词往往指称基本概念

功能性词语的意义往往很抽象，很难用更简单的英语词来解释（**paraphrase**），这些词指称的不是概念，而是联接概念的方式

摘自 *Jabberwocky*, by L. Carroll

这些无意义的词属于什么范畴？名词、形容词.....？属于词汇范畴还是功能范畴？

*'Twas brillig, and the slithy toves
Did gyre and gimble in the wabe;
All mimsy were the borogoves,
And the mome raths outgrabe.*

如何判断？

给词汇范畴下定义

例如，名词是什么？

名词可以指称几乎所有东西：*dog*, *dignity*,
necessity, *party*

词汇范畴作为等价类

更好是根据这些词的句法表现（**syntactic behavior**）将其定义为**等价类**（**equivalence classes**）

换言之，当我们说某些词语属于某个范畴X，我们想表达的是，这些词共有某种表现或分布，（**distribution**），而其他词都没有这种表现或分布

互相替换性

一个范畴中的诸元素可以互相替换
(**intersubstitutable**)；它们可以在合乎语法
(**grammaticality**)的前提下互换

互相替换性

可见，所有名词都拥有共同的特定分布；其中一种分布是名词在限定词（例如**the**或**some**）和句子末尾之间出现

- I saw the_____.
- John has some_____.
- There's a_____in my class.

上述结构都是名词的分布框架（**distribution frame**）

分布框架

每个分布框架都是一个测试

- There's a dog in my class.
- *There's a happy in my class.

*表示这个句子不合语法，即不是英语句子

dog是名词，happy不是

范畴与形态学

范畴也与形态学相关：词根（**base**）的范畴往往会影响附着于其上的词缀

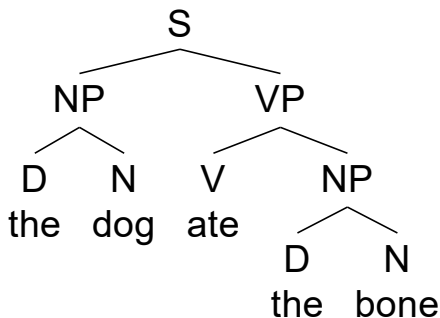
-ing 是动词后缀： *calling*, **talling*.

-ness 是形容词后缀： *tallness*, **callness*.

树状图

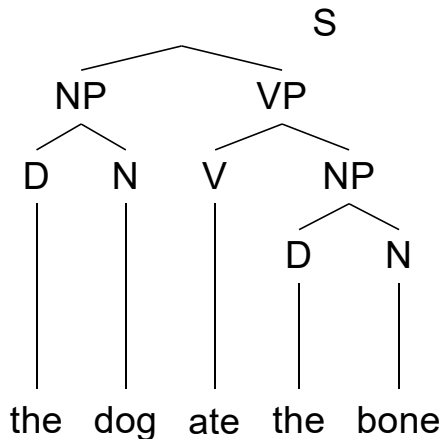
句子具有内在的分层结构，可以用树状图表示

- The dog ate the bone.



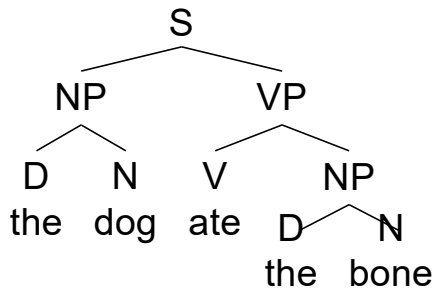
树状图

同一个树状图也可以表示如下：



组构成分

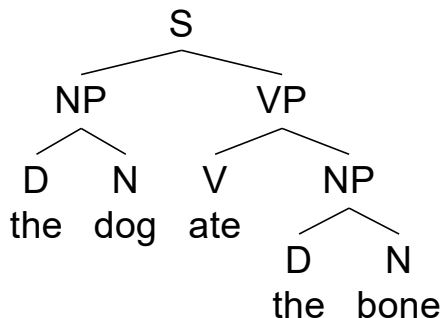
树状图表示的是词语被组合成短语
(**phrase**)，进而构成句子



树状图中的每个节点 (**node**) 都是一个组构成分
(**constituent**)；一些是短语 (包括S)，一些是词语

短语

和词语一样，短语也按照范畴划分



the dog 和 *the bone* 都是名词短语，而 *ate the bone* 是动词短语（VP）

短语的范畴

the dog 是**名词**短语，因为它可以被另一个名词替换（可替换性）

- The dog ate the bone.

- John ate the bone.

- The dog ate John.

类似地，*ate the bone* 是动词短语：

- The dog barked.

成分测试

可以通过**成分测试**（**constituency tests**）判断结构

用单个词语替换一个序列的词语就是一种成分组构测试

鉴于***the bone***可以被***John***替换，***the bone***是一个组构成分，而且是一个名词短语

问题

我们如何判断句子中的哪些词语能够形成组
构成分，哪些不能？

我们可以做一些成分测试（constituency tests）

常见的成分测试

- 替代法
 - 代词替代
 - “Do-So” 测试
- 独立法
- 并列法
- 位移法
 - 主题位移
 - 分裂句

代名化

有些词语是代形式（**pro-form**）：只要语境正确，它们能替换某个范畴中的组构成分，而不改变其意义

代词（*he, them, ...*）能代替名词短语（**pro-NP**）

- The dog ate the bone.
 - The dog was hungry. **He** ate the bone.
 - The bone disappeared. The dog ate **it**.

代名化

do so（和 *did so*, *does so*, ...）可以代替动词短语（pro-VP）：

- The dog ate the bone.
 - John ate the bone. The dog **did so**, too.
 - The dog was hungry. *Do so ate the bone.
 - John ate the bone. *The dog it, too.

非组构成分

不能通过成分组构测试意味着：这一词语序列不是属于某个范畴的组构成分（*the dog*不是VP），或者根本就不是什么组构成分：

- The dog ate the bone.
 - *The he the bone.
 - *The dog did so bone.

某一词语序列包含名词（N）并不意味着它是名词短语（NP）；某一词语序列包含动词（V）也不意味着它是动词短语（VP）

独立法：句子片段测试

一个组构成分可以作为对某个问题的简短回答

Paul ate at a really fancy restaurant yesterday.

Paul昨天做了什么？

- Ate at a really fancy restaurant.
- *Ate at.

Paul昨天在哪吃饭？

- At a really fancy restaurant.

并列法

并列连词**and, or, but**也可也用于成分测试

只有同一范畴的组构成分才可以并列连接

并列法

以下测试表明**the dog**是构成句子**The dog ate the bone**的名词短语成分：

- The dog and John ate the bone.
- The dog and the cat ate the bone.

以下测试说明了什么？

- The dog ate the bone and the fish.
- The dog ate the bone and barked.
- The dog ate the bone and drank the water.
- *The dog ate and cat ignored the bone.

位移法

另外还有位移法：如果我们能在一个句子中移动一个序列的词语，那么它们就是一个组构成分

有很多种位移测试，不过我们只关注主题位移法（**topicalization**）

主题位移作为一种成分测试

主题位移法指的是：只把某个组构成分移到句子的开头，而不改变句子的意义

- The dog ate the bone.
 - The bone, the dog ate.
 - *Ate the, the dog bone.
- I am terrified of those dogs.
 - Those dogs, I am terrified of.

分裂句

Paul ate at a really fancy restaurant yesterday.

■ Wh-分裂 (Wh-Cleft)

- What Paul did yesterday was to eat at a really fancy restaurant.
- Who ate at a really fancy restaurant yesterday was Paul.

■ It-分裂 (It-Cleft)

- It is at a really fancy restaurant that Paul ate yesterday.
- It is Paul who ate at a really fancy restaurant yesterday.

句法与词库

在我们的心灵词库中储存着各式各样的关于语素的信息

(/dɒg/ ; **N**; 一种四腿哺乳动物;)

(/run/ ; **V**; 以比步行更快的速度移动;)

当我们构建句子，句法会存取（**access**）**词汇范畴**的信息（**lexical category information**）

树形图与规则

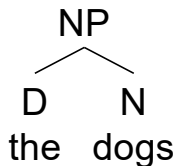
至此，我们看到句子具有树状结构

英语句法是一些概括性规则的集合，借以表明在英语中哪些树形图是可能存在的，即哪些句子类型是合乎语法的

我们把这些规则称为**短语结构规则**（**Phrase Structure Rules**）

短语结构规则 (PSR)

例如，我们见过，名词短语可能有如下结构：



对应的短语结构规则是：

$$\text{NP} \rightarrow \text{D N}$$

我们说这条规则生成了 (generate) 关于 the dogs 的上述树形图

中心语和修饰语

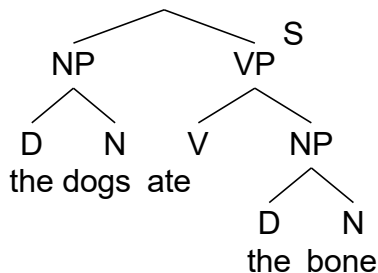
一般地，箭头左边的短语与箭头右边的一个语项（**item**）同属一个范畴。我们这个语项称为这个短语的**中心语**（**head**）

$$\text{NP} \rightarrow \text{D N}$$

根据这条规则以及由此规则生成的树形图，**N**是**NP**的中心语，箭头右边的其他部分（这个例子中是**D**）是**修饰语**（**modifier**）

其他短语结构规则

生成 *The dog ate the bone* 以及其他许多句子的短语结构规则的完全集（complete set）如下：

$$S \rightarrow NP VP$$
$$NP \rightarrow D N$$
$$VP \rightarrow V NP$$


每条规则分别生成这个树形图的不同部分

向心构式

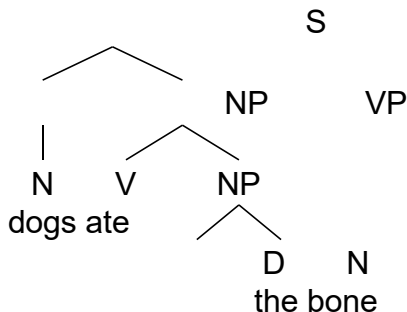
关于S的短语结构规则是不是有哪里不对劲？

包含中心词的短语结构规则是**向心的**
(**endocentric**)，不包含中心词的短语结
构规则是**离心的**(**exocentric**)

选择性的组构成分

以上的短语结构规则不能生成如下句子：

❑ Dogs ate the bone.



括号

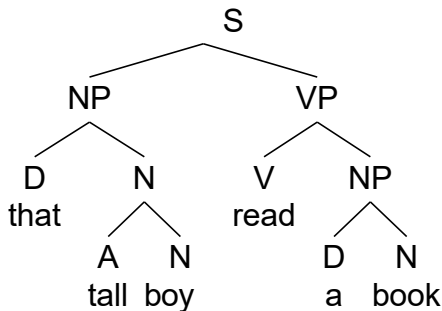
括号 (**parenthesis**) 表示某些修饰语是选择性/非强制性的:

$$\text{NP} \rightarrow (\text{D}) \text{N}$$

形容词

名词短语也可能包含形容词：

- That tall boy read a book.

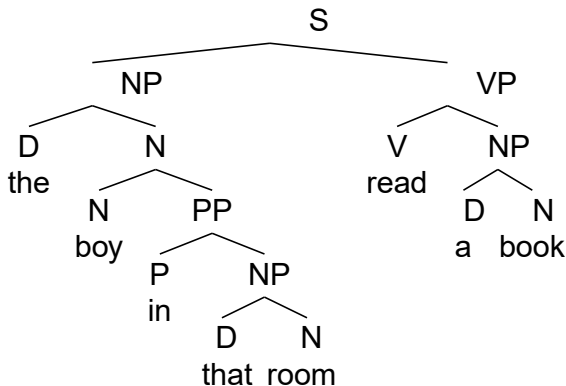


- NP \rightarrow (D) N
- N \rightarrow A N

介词短语 (PP)

介词短语 (**prepositional phrase**) 也可以修饰名词:

- The boy in that room read a book.



关于NP的修订版PSR

因此，我们需要给PP一条规则

$$\text{PP} \rightarrow \text{P NP}$$

我们也需要调整关于Ns的规则

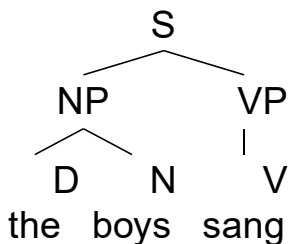
$$\text{N} \rightarrow \text{N PP}$$

这条修订过的规则可以生成我们之前不能生成的
大量新句子

动词短语

我们需要以类似的方式调整关于动词短语的规则

■ The boys sang.



修订后的规则：

■ $VP \rightarrow V(NP)$

前述PSR小结

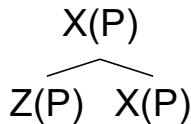
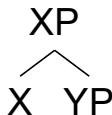
$$S \rightarrow NP VP$$
$$NP \rightarrow (D) N$$
$$N \rightarrow A N$$
$$N \rightarrow N PP$$
$$PP \rightarrow P NP$$
$$VP \rightarrow V (NP)$$

中心语、附接语和补语

中心语（**head**）决定短语所属的范畴

我们需要补语（**complement**）以使包含某些中心语的短语变得完整

可以有选择地添加附接语（**adjunct**），并不改变原有节点的标签（the label of the mother node）



附接语 vs. 补语

补语	附接语
一般地，每个短语只有一个补语	每个短语的附接语没有数量限制
为短语的完整是必要的	非必要的
附加于中心语	附加于中心语或一个短语

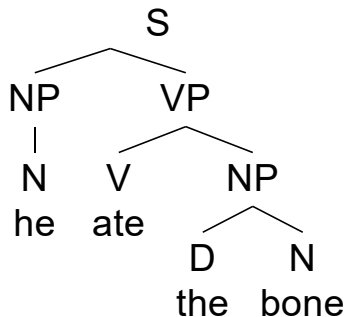
回到成分测试

成分测试其实就是一个句子，我们可以用这个句子来揭示另一个句子的结构

因此，这些测试也是带有结构的句子，这些结构是需要由语法生成的

代名化

例如，代词是一类特殊的名词，这点可以解释为什么代词可以代替名词短语（**pro-NPs**）



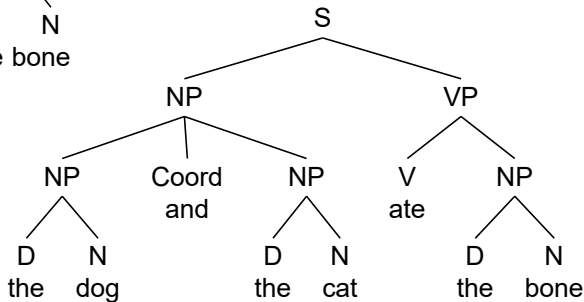
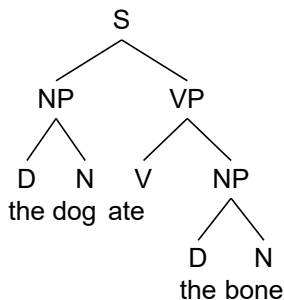
并列

我们需要新的短语结构规则以解释并列现象
(**coordination**) : $XP \rightarrow XP \text{ Coord } XP$

这条短语结构规则表明与并列相关的三点：

- 并列两边必须是**同一范畴**的组构成分
- 并列可以是**任意范畴**的组构成分的并列
- 生成的短语和其包含的并列成分属于同一范畴

包含并列的树形图



画树形图的一般步骤

两种方案：

- 自上而下
- 自下而上

自上而下的方案

1 画出基本形式



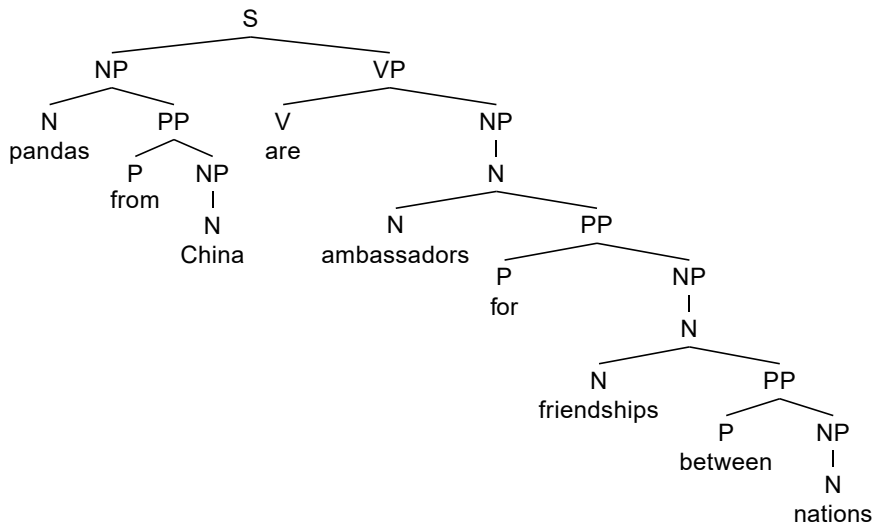
2 确定主语和谓语

3 向下逐步补充添加细节

用这种方法画出以下句子的树形图：

Pandas from China are ambassadors for friendships between nations.

Pandas from China



自下而上的方案

- 1 给句子中的每个词语标注词汇范畴 (**lexical category**)
- 2 确定哪两个词语组成一个单元
- 3 逐步向上构建整个句子