

作为语义—语用互动结果的语法形式^{*}

陆丙甫 南昌大学

提 要 本文通过语义性的“语义靠近象似性”和语用性的“可别度领先象似性”的互动去解释一系列跨语言的语序前后不对称现象。这里的前后，主要指核心词的前后。这些不对称包括语序变体、节律松紧程度和形态标志的数量多少等方面。例如在名词短语中，“前置（于核心名词的）定语”的语序相当稳定，变体较少，且跟核心名词的结合以及彼此间的结合较紧凑，并且形态较少（如汉语、英语）。相反，“后置定语”的语序相当自由，跟核心名词的结合以及彼此间的结合较松散，并且形态较丰富（如西班牙语等罗曼语言）。在动词短语中情况基本上相反。本文的分析导向两个结论：1）语义、语用上的表达功能决定了语序这一最基本的形式编码；2）语序又决定了节律和形态这两种形式编码。

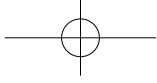
关键词 语义 语用 语序 节律 形态 因素互动

0 引言：人类信息处理能力限度与语序共性

胡裕树先生主编的《现代汉语》（第三版，1981：337）首次在国内强调了区分三个平面的重要性，指出必须区分三种语序：语义的、语用的和语法的^[1]，并举例说“我看你”“你看我”的区分是语义的；“你哥哥来了吗？”“来了吗，你哥哥？”的区分是语用的；而“客来了”“来客了”的区分是语法的，因为从主谓结构变成了动宾关系。笔者多年来根据三个平面思想，对语序规律以及语序、节律、形态之间的相关性，做了一系列研究。本文对这些研究做一个小结。

语言是人类最重要的信息载体。人类在线信息处理的能力是有限的。这主要表现为短时记忆和注意力的限度为“七块左右（ 7 ± 2 ）”（Miller 1965a）。

^{*} 本文内容，曾在“纪念胡裕树先生诞辰一百周年学术讨论会”上报告（2018/8/26-27，复旦大学），受益于与会者的反馈。本文的写作，也得益于卢英顺、徐阳春、李占炳、叶爱等先生的反馈、交流。尚存问题均归笔者。谨以此文纪念笔者敬爱的导师胡裕树先生。



这里的“块（chunk）”，即信息块。有趣的是，首先指出这个限度的 Miller 自己竟不知道语言结构中的信息块该如何定义（Miller, 1956b）：

我们通常能够把一个 20 个词组成的句子听过一遍之后加以复述。这句话包含多少单位？100 个字母？30 个音节？20 个词？6 个短语？2 个分句？抑或 1 个句子？我们可以知道它包含了大约 120 比特的信息，因为比特的定义独立于我们对句子的主观组织。但是板块的本质是它是人为的。例如，一个除了字母外对英语一无所知的人 would 认为这句话里有 100 个单位；但一个英语很好的人 would 认为这句话里只有 6 个单位。我们不能离开了听者去定义组织的单位。

Chomsky（1965：3）对语言的看法纠正了 Miller 的失误。

语言学主要研究在某个单纯的话语社区的理想的说话—听话者，他充分掌握了这种语言，并且不受影响于那些跟语法无关的因素，如记忆限制、分心、兴趣转移以及其他种种在具体运用中发生的失误（包括偶然性的或特色性的）。

我们研究英语显然不能以“除了字母外对英语一无所知”的人的语感为依据。而一个英语很好的人，即英语母语者，根据 Miller 的天才猜测，可以把一个足够复杂的英语句子分成六个短语。但问题是短语具有递归性，大短语中可包含小短语，如句子本身可看作一个最大的短语，主、谓两部分也可各自看作是个短语，并且 Miller 所说的六个短语，其中每个内部都可以包含更小的短语。如何明确定义 Miller 所说的这些短语呢？

定义短语离不开“核心词”。从内部看，根据 X' 理论，每个短语都由一个核心词投射而来^[2]。从外部看，短语也需要一个核心词去控制它。因此，构成 Miller 假想的这个英语句子的信息块，应该是六个短语加上一个动词，即是七个信息块，例如：

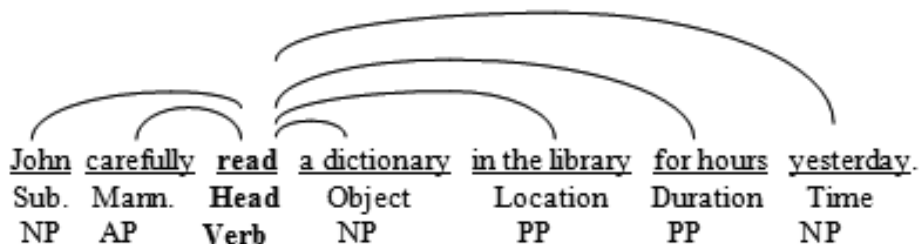


图 1 核心词及其从属语

这个句子由六个短语组成，这一点上满足了 Miller 假想中那个英语句子的基本结构。但是 Miller 假想的句子里有 100 个字母，20 个词，但是我们这个句子中只有 57 个字母，11 个词，比起 Miller 的假想句子简单得多。要把这个句子凑满 100 个字母、20 个词是很容易的，因为其中任何一个短语都有无限的扩展可能。但是我们可以注意到，这个句子要再增加一个直接跟核心动词有关的短语，是极为困难的。

一个句子由不超过七个左右的结构成分构成，也许这就是人类语言最基本的语序共性。这个共性反映了人类语言既充分利用，并且也受制于人类一般的信息处理能力。

以上得出的语言结构中的信息块，可以称为一个语言结构体的“直接成分”（direct chunk 或 direct constituent, DC）^[3]。事实上也就是语序研究中的语序单位，不过是以往无法对此给出一个明确定义而已。之所以如此，是因为学界深受句法研究中传统的两分法——树形图的束缚。若把语言想象成天体那样的轨层结构，则整个结构图景就大大简化了。

我们前面说到短语的一个特征是可以无限递归。要克服无限递归，找出每个结构模式中的有限组成成分，就要找到无限递归的根源。无限递归的原因是核心的不断转移。把某个结构中的任何一个从属语中的某个成分，当作一个核心，就能带上自己的从属语而引进一个新的短语。因此，只要锁定一个核心，就能使成分的分析成为有限的。

回过头来看，Chomsky 虽然纠正了 Miller 的失误，但同时也犯了另一个错误，那就是把语言交际过程中无时无刻不存在的，绝非偶然性的短时记忆和注意力的限制，看作跟语言机制无关的因素。这个时刻限制着语言形式的重要因素，决定了人类语言最基本的共性，即信息块在数量上的普遍限制。

1 语义靠近象似性（SPI）

这方面，语言结构类似于宏观或微观的物理世界。太阳是核心，八大行星在不同的轨道中围着它转，构成一个太阳系结构。原子也由若干电子围绕着一个原子核而构成，并且也是原子核决定了原子的基本性质。

我们可通过“向心切分法”（每次两分后都选择核心部分继续切分下去）将树形图加以简化，结果得到“主干成分树形图”。

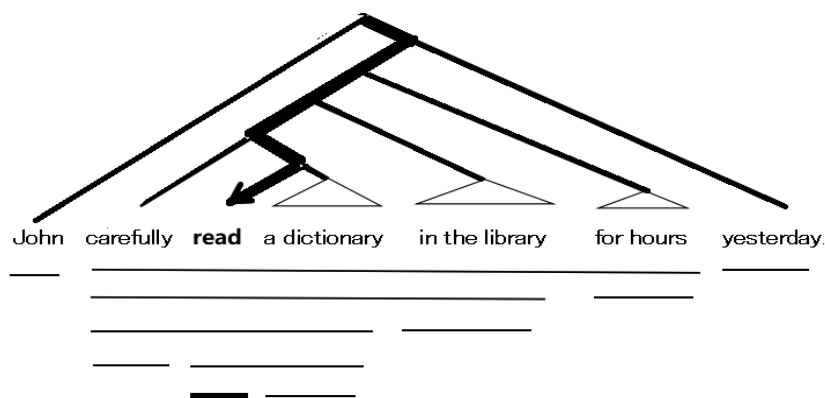
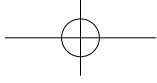


图 2 主干树形图

“向心切分法”可看作直接构成句子的信息块，即 DC 的操作性定义。上图中通向核心动词 **read** 的粗黑线条可看作简化树形图的躯干，箭头也可代表整个结构的生长点。跟躯干直接连接的较粗的线条是“主干”，三角形表示必须排除的枝枝叶叶。

传统树形图中是不同的结构模式成分交叉纠缠在一起，因此曾经被批评为“叠床架屋”。比起传统树形图，主干树形图的好处是明确地把共存于一个结构体中不同的结构模式做了明确区分：一个核心决定一个主干树形，一个主干树形代表一个结构模式：核心词加上其所有主干成分，构成一个结构模式。为了进一步凸显核心的关键作用，还可以进一步把主干树形转化成“向心轨层”结构：

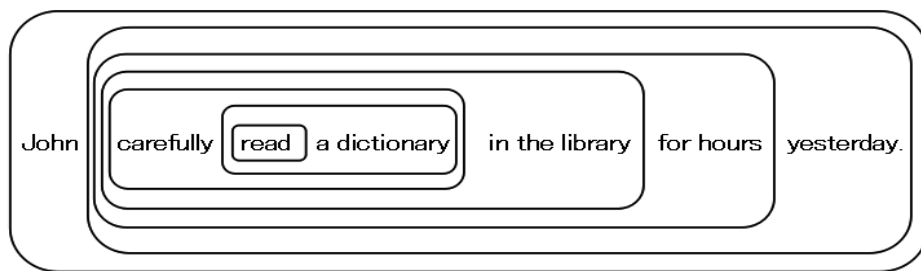


图 3 轨层图

如同 X 杠杆理论所体现的，极端重视核心词的作用，认为一个短语是一个核心词性质的投射。我们不过是进一步在分析一个短语时，把一个结构体

中由不同核心词投射而成的，即不同轨层的短语明确区分了开来。

上述语言结构图景简化的效果，从下面例子中可以看出。



图4 动词和五个状语的语序

- (1) a. [他[上星期[[在实验室[用计算机[连续地[工作了]]]]三天]]]
 b. [He [[[[[worked] continually] with computers] in the lab] for three days] last week]]

为了简化问题，尽量减少干扰因素，上图去掉了主语，因为主语带有极大的话题性这一语用干扰因素^[4]。剩下六个成分“时位(T)、时量(D)、处所(L)、工具(I)、方式(M)”和动词核心，数学上可能的排列应该有 $6 \times 5 \times 4 \times 3 \times 2 = 720$ 种。

但是我们目前调查到的作为某种语言中的无标记语序只有上图的六种。

根据这六种语序，可以建构起上图六个成分组成的垂直轨层序列，它反映了五个从属语跟核心动词的距离远近。由于这个轨层距离顺序的限制，每个从属语都只能在自己的轨层中落实两个位置，理论上可能的排列就下降到 $2 \times 2 \times 2 \times 2 \times 2 = 32$ 种（作为某种语言中无标记的基本语序）。

实际上可能的排列要少得多，我们目前只发现以上6种。这是因为除了轨层的限制外，还受到其他一些限制，其中之一就是“如果时位T和时量D分置动词两旁，只能是时位在前而时量在后”，（即 $[T \cdots V \cdots D]$ ，如汉语的“上星期…工作了三天”）。这样，包含 $[D \cdots V \cdots T]$ 的8个序列就从32种语序中被排除了。排除它们的原理就是“整体—部分”的优势顺序（陆丙甫，2010）。这样，剩下的理论上可能的基本语序就只有24种了。但这仍然是实际发现的