语言描写的三个模型

N Chomsky 著 张和友译 冯志伟校

冯志伟的说明

《语言描写的三个模型》是 Chomsky 的早期著作，于 1956 年发表在《信息论杂志》上，Chomsky 在语言学的历史上首次采用 Markov 模型来描写自然语言，对于有限状态模型、短语结构模型和转换模型等三个模型，从语言学和数学的角度进行了理论上的分析，建立了形式语言理论，具有划时代意义。Chomsky 在这篇文章中明确地提出：“语言理论试图解释说话人在其有限的语言经验的基础上生成和理解新的句子并拒绝其他不合语法的新序列的能力。”这一直是生成语法追求的目标，也是最简方案的语言学理论所追求的目标。后来 Chomsky 在 1957 年发表的《句法结构》实际上是这篇文章的通俗解释，1995 年发表的《语言学理论的最简方案》实际上是这篇文章思想的进一步发展，要深入理解 Chomsky 的形式语言理论和最简方案理论的实质，应该首先从《语言描写的三个模型》入手。这篇文章的英文发表于《信息论杂志》，属于理科杂志，国内很少有语言学家读过这篇文章，因此，我们把这篇重要的文章翻译成中文，可以醒人之耳目，也可以弥补读者查找资料的困难。

摘要

我们研究了好几种语言结构的观念，看看它们能否提供简单而显露的语法来产生英语中的所有句子，并且仅仅产生这些句子。我们发现，产生由一种状态转移到另一种状态符号的有限状态马尔可夫过程，没有哪一种过程能够充当英语语法。并且，从统计的角度看，产生接近于英语（句子）的 n 元组的这些过程的特殊次类，随着 n 的增大，并不是越来越接近于跟一部英语语法的输出相匹配。我们将“短语结构”的概念形式化，并且证明“短语结构”给我们提供了一种描写语言的方法，这种方法本质上是很有力的，不过仍然可以描写成有限状态过程的一种相当基本的类型。然而，只有当将范围限制在简单句的较小子集中时，这一方法才是成功的。我们研究了一套语法转换的形式属性，这些转换将带有短语结构的句子变成带有导出短语结构的句子，这表明转换语法是具有跟短语结构语法相同基本类型的过程；并且如果短语结构描写限制在所有其他句子通过反复运用转换得以建构的核心（kernel）的简单句上，英语的语法将大大简化。这也表明，这种关于语言结构的观点提供了某种使用和理解语言的悟性。

1. 导言

在语言的描写研究中有两个中心问题。语言学家关注的一个基本问题是找到自然语言的简易而显露的语法。同时，通过研究这种成功语法的属性和澄清作为这种语法基础的基本观念，语言学家希望找到语言结构的一般理论。我们将检查这些相关研究的某些特性。

一种语言的语法可以看作这种语言结构的理论。任何科学理论都是建立在一定的有限观察的基础上；通过建立依照一定的假说结构表达的普遍法则，科学理论试图解释这些观察，试图显示这些观察是如何相关的，进而试图预测无限多的新现象。数学理论还有一个特征，即预测是从理论本身精确的得出的。类似地，一种语法是建立在对于有限句子的观察之上（语

言学家的语料库），并通过建立普遍法则（语法规则）将观察到的有限句子的集合投射到（project）包含无限的合乎语法的句子集合中。普遍法则是按照所分析的语言中诸如特殊的音位、词、短语等等之类的假说结构来构造的。一种形式化合适的语法应该无歧义的决定合乎语法句子的集合。

普通语言学理论可以看作一种元理论（meta-theory），它关心的是这样的问题：就每一种具体语言而言，在有限句子的语料库的基础上，如何去选择那样一种语法。特别是，这种理论要考虑并试图阐述合法的句子集合与所观察的句子集合之间的关系。换句话说，语言理论试图解释说话人在其有限的语言经验的基础上生成和理解新的句子并拒绝其他不合语法的新序列的能力。

假定对于许多语言而言，存在一定的合法句子的纯粹事实与一定的不法序列的纯粹事实，分别如英语中（1）、（2）这样的句子：

（1）John ate a sandwich.

（2）Sandwich a ate John.

在这种条件下，我们验证一种议定的语言理论适切性（adequacy）的办法是看看对于每一种语言来说，一定的纯粹事实能否用依照这一理论构建的语法加以合理的解决。比方说，如果英语的大规模语料库中碰巧没有包含（1）或者（2），那么，我们就问：为这一语料库设计的语法能否将语料库设计成包括（1）而排除（2）的样子。即使这些纯粹的事实为已知的孤立考虑的语言的语法的适切性仅仅提供微弱的验证，它们却为任何一般语言理论以及由语言理论产生的语法集合提供了有力的验证，既然我们主张就每种语言而言，纯粹事实应该以一种固定的和预定的方式加以适当的解决。我们可以采取一定的步骤来构建“合法句子”的操作特征，这将给我们提供为有效确定语言学任务所需求的纯粹事实。例如，我们观察到，说英语的人用语料库中句子的正常语调来朗读（1），而对（2）中的每个词却用降调来读，像诵读任何没有关系的词构成的序列。其他同类的区别性准则也能够加以描写。

在希望为观察到的句子与合法句子之间的一般关系提供满意的解释之前，我们必须更多的了解每一个集合的形式特征。本文关注的是合法句子集合的形式结构。我们将范围限制在英语，并且假定具有英语句子和非句子的直觉知识。那么，我们可以问：对于以一种有意义和令人满意的方式描写英语句子集合的英语语法，我们需要什么样的语言理论作为基础呢？

对一种语言进行语言分析的第一步就是为这种语言的句子提供有限的表征系统。我们假定这一步已经完成，并且我们只用语音或者字母记录来处理语言。这样，我们所说的一种语言，是指句子的集合（有限的或者无限的）。每个有限长度的句子都是由有限的字母符号构成的。如果 A 是一个字母表，我们说通过联结 A 的符号所形成的任何形式是 A 中的一个符号串（string）。我们所说的一种语言 L 的语法，是指某种装置，这种装置产生属于 L 中的句子的一切符号串，并且仅产生这些符号串。

无论我们最终决定如何去构建语言理论，都将必然要求任何语言的语法一定是有限的。于是，任何语言理论都仅使语法的有限集合变得可用；因此，在一般意义上讲，按照由任何具体理论提供的语言结构的观念，无法估计的多数语言严格来说是不能描写的。假定给定一种语言结构的理论，那么，问下面这样的问题总是适宜的：

（3）绝对处在所给定的描写类型之外的有意义的语言存在吗？

特别是，我们应该问问英语是否是这样的语言。如果是的话，那么原来给定的语言结构的概念必定是不适当的。假如对于（3）的回答是否定的，那么我们将接着问下面这样的问题：

（4）我们能够为一切有意义的语言合理的构建简明的语法吗？

（5）这种语法在下列意义上是显露的吗？即：这种语法所呈现的句法结构能够支持语义分析，能够为语言的使用和理解提供洞察力，等等。

我们先按照描写的可能性与复杂性（问题（3）、（4））来检验一下关于语言结构的各种观念。然后，在第六部分，我们按照问题（5）简要的考虑一下同样的理论；我们将看到，为了语言学的目的，我们会独自得出关于相对适切性的相同结论。

2 有限状态马尔可夫过程

2.1 用有限装置生成无限句子的最基本的语法是一种建立在跟特别简单的信息源相似的语言观念基础上的语法，这种特别简单的信息源就是有限状态马尔可夫过程（finite-state Markov process）[①] 。具体说，我们将有限状态语法 G 定义为一个包含有限状态的系统：$S_0\cdots,S_q$，转换符号的集合 A=｛a_{ijk}丨0≤i，j≤q；1≤k≤N_{ij}｝，据说是相联的语法的某些配对状态的集合 C=｛(S_i，S_j)｝。随着系统从状态 S_i 转移到 S_j，产生符号 a_{ijk}，a_{ijk}∈A。假定：

（6）$S_{\alpha 1}\cdots$，$S_{\alpha m}$

是$\alpha_1=\alpha_m=0$ 时语法 G 的状态的一个序列，且（$S_{\alpha i}$，$S_{\alpha i+1}$）∈C （i<m）。随着系统从状态 $S_{\alpha i}$ 转移到 $S_{\alpha i+1}$,产生符号（7）：

（7）$a_{\alpha i\alpha i+1K}$

K≤$N_{\alpha i\alpha i+1}$ 。我们用弧形∩表示毗连（concatenation）[②]，我们认为，序列（6）产生一切形如（8）的句子：

（8）$a_{\alpha 1\alpha 2K1}\cap a_{\alpha 2\alpha 3K2}\cap\cdots\cap a_{\alpha m-1\alpha mKm-1}$

对于 K_i 的所有适宜选择（即对于 K_i≤$N_{\alpha i\alpha i+1}$），包含并且仅包含上述句子的语言 L_G 称为由语法 G 产生的语言。

于是，为了产生 L_G 的一个句子，我们将系统 G 置于初始状态 S_0 中，伴随着从一个状态转移到下一个状态，产生一个 A 的相关转移符号的过程，我们经历一系列的相关状态，最后仍以 S_0 结束。如果语言 L 是通过某种有限状态语法 G 产生的句子的集合，那么我们称 L 是有限状态语言。

2.2 我们不妨将转换符号的集合 A 看作英语音位的集合。我们能够试图构建一种有限状态语法 G，这种语法将生成每一个属于英语的合法句子的音位符号串，并且只生成这些符号串。我们马上看到，如果将 A 看成英语的语素[③]或者词的集合，并构建语法 G，结果语法 G 恰好产生这些单位的合法串，那么，为英语构建有限状态语法的任务可以大大简化。这样，通过有限的规则集合，就完成了这部语法。这些规则给出每一个单词或者语素依据其出现环境的音位拼读形式。我们将在 4.1 和 5.3 中简要考虑这些规则的情形。

在直接探究为英语的语素或者词的序列构建有限状态语法这一问题之前，我们来调查一下有限状态语言的绝对限制。假定 A 是语言 L 的字母表，a_1 ……a_n 是这个字母表的符号，并且 S=$a_1\cap\cdots a_n$ 是 L 的一个句子。我们认为，对 L 而言 S 具有一个（i, j）依存的充要条件是：

（9）（i）1≤i≤j≤n

（ii）存在符号 b_i、b_j∈A，具有如下属性：S_1 不是 L 的句子，S_2 是 L 的句子，其中 S_1 是通过用 b_i 替换 S 的第 i 个符号（即 a_i）从 S 形成，S_2 是用 b_j 替换 S_1 的第 j 个符号（即 a_j）从 S_1 形成。

换句话说，如果用 b_i 对 S 的第 i 个符号 a_i 的替换（$b_i\neq a_i$）要求相应的 b_j 对 S 的第 j 个符号 a_j 的替换作为结果符号串属于 L（$b_j\neq a_j$），那么，对 L 而言 S 具有一个（i, j）依存。

我们认为 D=｛（α_1, β_1），…（α_m, β_m）｝对 L 中的句子 S 而言是一个依赖集的充要条件是：

（10）（i）当 1≤i≤m 时，对于 L 来说，S 有一个（α_i, β_i）依存

[①] 参文献[7]。有限状态语法可以按照状态图用图表的方式加以表达，如文献[7]的第 15 页脚注。

[②] 参文献[6]，附录 2 是一个线性代数的公理化。

[③] 我们的语素指的是语言中最小的语法功能成分，如“boy”、“run”、“running”中的“ing”、“books”中的“s”等。

（ii）对于每个 i, j 而言，$\alpha_i \leqslant \beta_j$

（iii）对于每个 i, j 而言，$i \neq j$，$\alpha_i \neq \alpha_j$，$\beta_i \neq \beta_j$

这样，对 L 中的 S 而言，在一个依赖集中，每两个依赖集在两个项（item）上各不相同，S 中的每个决定成分先于被决定成分，其中，我们将 $a_{\alpha i}$ 描述成决定 $a_{\beta i}$ 的 选择。

很显然，如果在 L 中，S 具有一个 m 项依赖集，那么，在生成语言 L 的有限状态语法中，至少有 2^m 种状态是必需的。

这种观察使我们能够确定有限状态语言的必要条件。

（11）假定 L 是一种有限状态语言。那么，存在 m，L 中没有一个句子具有多于 m 的依赖集合。

脑子里记住这个条件，我们能很容易构建许多非有限状态语言。比方说，（12）中描述的语言 L_1、L_2、 L_3 不能用任何有限状态语法加以描述。

（12）（i）L_1 包含 a⌒b，a⌒a⌒b⌒b，a⌒a⌒a⌒b⌒b⌒b，……，一般地，L_1 包含所有这样的句子：这些句子含有 n 个 a，其后恰好跟着 n 个 b，并且 L_1 只含有这些句子；

（ii）L_2 包含 a⌒a，b⌒b，a⌒b⌒b⌒a，b⌒a⌒a⌒b，a⌒a⌒b⌒b⌒a⌒a，……，一般地，L_2 包含所有的“镜像”句子，这些句子含有符号串 X，其后跟 X 的倒置排列符号串，并且 L_2 仅仅含有这些句子；

（iii）L_3 包含 a⌒a，b⌒b，a⌒b⌒a⌒b，b⌒a⌒b⌒a，a⌒a⌒b⌒a⌒a⌒b，……，一般地，L_3 包含所有这样的句子：符号串 X 之后跟着相同的符号串 X，并且 L_3 仅仅包含这些句子。

比方说，在 L_2 中，对于任何 m，我们都能够找到一个句子具有依赖集 D={（1，2m），（2，2m－1），…，（m，m+1）}[④]。

2.3 现在转到英语上来。我们发现，存在无穷的句子集合，它们具有包含超过任何固定项目（term）的依赖集。比方说，设定 S_1、S_2，…是陈述句，那么，下面这些都是英语的句子：

（13）（i）if S_1，then S_2。

（ii）either S_3，or S_4。

（iii）The man who said that S_5, is arriving today.

这些句子具有“if”－“then”，“either”－“or”，“man”－“is”这样的依赖集。不过我们可以像（13i）、（13ii）、或（13iii）自身那样选择出现于相互依赖词之间的 S_1、 S_3、S_5。继续以这种方式构建句子，我们就得到英语的次类，这些次类恰好具有（12）中的语言 L_1 和 L_2 的镜像属性。因此，英语不满足条件（11）。英语不是有限状态语言，因之我们被迫放弃所讨论的语言理论，因为它不满足条件（3）。

我们可以通过主观规定英语的句子长度存在有限上界来避免这种结果。然而这种做法无益于有用的目的。重要的问题是，存在着这种基本的语言模型从本质上讲不能加以处理的句子构造过程。如果对这些过程的操作不进行有限限制，那么，我们可以从文字上证明这种模型的不适宜性。如果对这些过程加以限制，那么，严格说来，构造有限状态语法将不是不可能的（因为一个清单就是一个小的有限状态语法），但是这种语法如此复杂以致没什么作用或意义。下面，我们将研究一个能处理镜像语言的语法模型。这种模型在无限状态下的额外力量由下列事实得到反映：如果作上限限制，那么这种模型更为有用和具有外显性。一般说来，做出语言是无限的这样的假设，目的在于简化描写[⑤]。语法如果没有递归步骤（即上面讨论过的模型中的封闭环）就会变得极其复杂——实际上，我们将证实，这种语法对自然语

[④] 在 L_1 的情况下，（9ii）的 b_j 可以看作一个同样的成分 U，具有如下属性：对于所有 X 而言，U⌒X= X⌒U=X 。那么，D_m 对 L_1 中长度为 2m 的句子来说也是一个依赖集。

[⑤]注意，一种语法必须反映和解释说话人产生和理解新句子的能力，这些新句子可以比他以前听过的任何句子都长得多。

言来讲比符号串或语素类序列的清单好不了多少。如果语法的确有递归装置，它将会产生无限多的句子。

2.3 尽管我们看到没有哪一种自左至右产生句子的有限状态马尔可夫过程能够充任英语语法，我们仍可以研究构建具有如下装置（device）的序列的可能性，这种装置以某种并非不重要的方式跟令人满意的英语语法的输出越来越紧密匹配。比方说，假定对于固定的 n 来说，我们按照下列方式构建有限状态语法：这种语法的一种状态跟每个长度为 n 的英语词的序列相联，并且当系统处于状态 S_i 时，产生词 x 的概率跟 x 的条件概率（conditional probability）相等（给定界定 S_i 的 n 个词的序列）。这个语法的输出通常被称作英语的第 n+1 个近似值（approximation）。很明显，随着 n 的增大，这种语法的输出将变得越来越像英语，因为序列越长，直接从英语样品中产生出来的概率就越高，概率是由英语样品决定的。这个事实有时会使我们认为，语言结构理论建立在这种模型之上。

不论对于这个意义上的统计学上的近似值的兴趣是什么，明显的情况是，这种统计上的近似值都没有阐明语法的问题。在一个符号串（string）（或是其组成成分）的频度（frequency）和它的合法性（grammaticalness）之间不存在一般关系。为了能更清楚地明白这一点，我们来看（14）这样的符号串：

（14）colorless green ideas sleep furiously

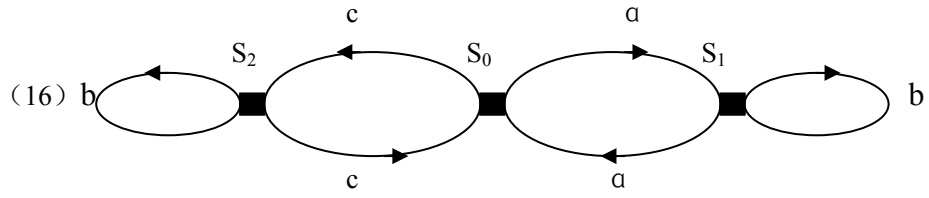
（14）是一个合法的句子，尽管假定（14）中的词在以前未曾配对出现是合理的。注意，一个英语的说话人会用英语句子的通常语调模式来理解（14），可是他也同样会用在每个单词上用降调的办法来理解不熟悉的符号串（15）：

（15）furiously sleep ideas green colorless

这样，（14）跟（15）之间的差别正如（1）跟（2）之间的差别；我们暂时的对于合法性的操作准则支持我们的直觉，即：（14）是一个合法的句子而（15）不是。我们可以部分程度上将语法的问题表述为解释与重构英语说话人承认（1）与（14）是合法的而拒绝（2）与（15）是合法的这一能力。但是没有近似值模型的规则能够区分（14）和（15）（或者区分数目不定的类似配对儿）。随着 n 的增大，英语的第 n 个近似值将排除（由于越来越未必发生）一度增长的合法句子的数目，而另一方面它仍然包含数目庞大的彻底不合法的符号串[⑥]。我们不得不得出如下结论：在这方面显然不存在解决语法问题的有效方法。

我们注意到，尽管对于每个 n 来说，一个 n 阶近似值的过程可以表达成一个有限状态马尔可夫过程，但是反之则不然。例如，将带有（S_0， S_1），（S_1，S_1），（S_1， S_0），（S_0，S_2），（S_2，S_2），（S_2， S_0）三种状态的（three-state）过程看成其唯一的联结状态，将带有 a，b，a，c，b，c 三种状态的过程看成各自的转换符号。这一过程可以用下列状态图表示：

（16）



这一过程能够产生如下句子：a⌢a, a⌢b⌢a, a⌢b⌢b⌢a, a⌢b⌢b⌢b⌢a,…，c⌢c, c⌢b⌢c, c⌢b⌢b⌢c, c⌢b⌢b⌢b⌢c,…，但是不能产生 a⌢b⌢b⌢c, c⌢b⌢b⌢a 等这样的句子。所生成的语言的句子具有任何有限长度的附属物。

在§2.4 中，我们提出理由说在近似规则与合法性之间不存在有效的相互联系。从接近英

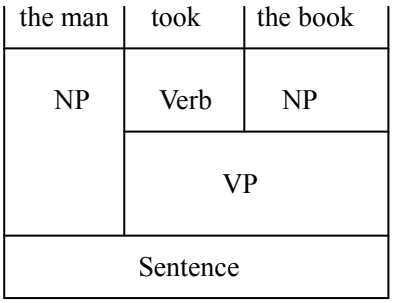
[⑥]这样我们总是看到 n+1 个词的序列，其中前面的 n 个词和后面的 n 个词都可以出现，但不可以在同一句子中出现（比方说，将（13iii）中的“is”换成“are”，并且选择具有任一合要求长度的 S_5）。

语的规则的角度看，如果我们给定具有已知长度的符号串，我们将看到，整个清单从上而下所散布的既有合法的句子，又有不合法的句子。因此，统计学上近似的观念看起来跟语法无关。在§2.3 中我们指出，一个范围更广的过程，即所有产生转换符号的有限状态马尔可夫过程都不包括英语语法。也就是说，如果我们构建一种只产生英语句子的有限状态语法，我们知道这种语法不能产生无限数目的这些句子；特别是，它不能产生无限数目的真句子、假句子以及能够明了提出的合乎情理的问题与类似问题。下面，我们将考察更广范围的过程，这些过程可能为我们提供英语语法。

3 短语结构语法

3.1 通常情况下，句法描写是按照所谓的“直接成分分析”（immediate constituent analysis）给定的。在这类描写中，句子的词组合成短语，这些短语被分成更小的成分短语，如此等等，直至达到最后的成分（一般来说就是语素[3]）。这些短语然后分成名词短语（NP）、动词短语（VP），等等。例如，（17）可以随同附表分析如下：

（17）



明显地，跟词对词（word-by-word）模型相比，按照这种方式描写句子准许相当大的简洁性，因为由像 NP 这样的复杂类型的短语构成的复合结构在语法中可以仅仅表达一次，并且这种结构可以在句子结构的各个地方作为一个建构片儿（building block）使用。我们现在要问的是：什么样的语法形式跟语言结构的这种概念相对应？

3.2 我们将一种短语结构语法定义为一个有限词汇（字母表）V_p，V_p 中初始符号串的有限集合∑和一个形式规则的有限集合 F：X→Y,其中 X 与 Y 是 V_p 上的符号串。每一条这样的规则都可以解释为如下指令：将 X 重写为 Y。鉴于即将直接显露的原因，我们要求在每一个[∑, F]语法中

（18）$\sum$：$\sum_1 \cdots, \sum_n$

F：$X_1 \rightarrow Y_1$

$\vdots$

$X_m \rightarrow Y_m$

通过用某种符号串来替换单一的符号 X_i 可以从 X_i 形成 Y_i。被替换的符号和用来替换的符号串都不可能是注释 4 中的成分 U 本身。

给定[∑, F]语法（18），我们认为：

（19）（i）符号串 ß 从符号串 α 导出，如果 $\alpha = Z \frown X_i \frown W$, $ß = Z \frown Y_i \frown W$, $i \leqslant m$[⑦]

（ii）符号串 S_t 的一个推导式是一个由符号串构成的序列 $D=(S_1, \cdots, S_t)$，其中 $S_1 \in \sum$，并且对于每个 $i < t$ 来说，S_{i+1} 都可以从 S_i 导出；

（iii）如果按照（18）存在一个 S 的推导式，那么符号串 S 可以从（18）推导出来。

---

⑦ Z 或 W 在这种情况下可能是成分 U 本身（比较注释 4）。注意，既然我对（18）加以限制以便排除将 U 内含地计算在规则 Γ 的右侧或左侧；并且既然我们要求只有左侧的一个单一符号可以在任何规则中被替换，那么可以得出结论说：Y_i 必须至少跟 X_i 等长。这样，我们对于在（19iii）、（19V）意义上可推导性与终端性（terminality）有一个简单的决定程序。

（iv）如果不存在从 S_t 导出的符号串，那么，S_t 的推导终止；

（v）如果符号串 S_t 是一个终止的推导式的最后一列，那么它就是一个终极符号串。

于是一个推导式大约跟一个证明类似，将∑看作公理系统，将 F 看作推理规则。如果 L 是符号串的集合，可以从某种[∑, F]语法推导而来，那么，我们说 L 是可推导语言；同样，如果 L 是终极符号串的集合，可以从某种[∑, F]系统推导出来，那么，我们说 L 是终极语言（terminal language）。

在每一种有意义的情况下，都存在一个终极词汇 V_T（$V_T \subset V_P$），V_T 恰好描述终极符号串；这是在下列意义上说的：每个终极符号串是 V_T 中的符号串，并且 V_T 中没有哪个符号按照 F 的规则中发（？）任何一个进行重写（rewritten）。在这种情况下，我们可以将终极符号串解释成构成所分析的语言（以 V_T 为其词汇），将这些符号串的推导解释成它们的短语结构。

3.3 作为形式系统（18）的一个例子，请看下面一小部分英语语法：

（20）∑：#∩Sentence∩#

F：Sentence→NP∩VP

VP→Verb∩NP

NP→the∩man, the∩book

Verb→took

在源自（20）的推导式中，我们特别有：

（21）D_1：#∩Sentence∩#

#∩NP∩VP∩#

#∩NP∩Verb∩NP∩#

#∩the∩man∩Verb∩NP∩#

#∩the∩man∩Verb∩the∩book∩#

#∩the∩man∩took∩the∩book∩#

D_2：

#∩Sentence∩#

#∩NP∩VP∩#

#∩the∩man∩VP∩#

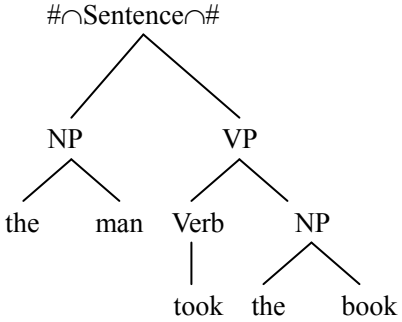
#∩the∩man∩Verb∩NP∩#

#∩the∩man∩took∩NP∩#

#∩the∩man∩took∩the∩book∩#

这些推导式显然是等同的；它们的不同仅仅在于规则的使用的顺序。我们可以通过显明的方式构建跟推导式相应的图表对这种等同加以图示。D_1 和 D_2 都可以归结为如下图表：

（22）



图（22）跟表（17）一样给出了终端句子“the man took the book”的短语结构。一般地，给定一个符号串 S 的推导式 D，我们认为，如果在跟 D 对应的图表中，S 的子串 s 可以回溯

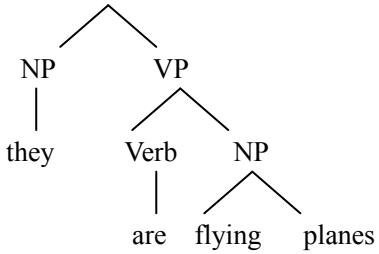
为一个单一的节点（node），那么，s 就是一个 X，并且将这个节点标为 X。例如，给定跟（22）对应的 D_1 或 D_2，我们认为，“the ∩man”是一个 NP，“took∩the∩book”是一个 VP，“the∩book”是一个 NP，“the ∩man∩took∩the∩book”是一个句子。但是，“man∩took”绝对不是这个符号串的一个短语，因为它不能回溯为任何一个节点。

当我们试图为英语构建最简单的可能语法[∑,F]时，我们看到，某些句子自动接受非等同的推导式。跟（20）相伴随，英语语法必将包含如下规则：

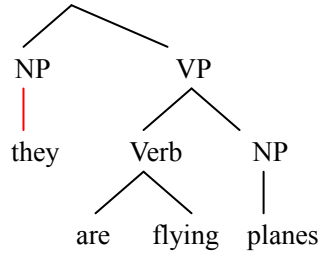
（23）Verb→are∩flying
Verb→are
NP →they
NP →planes
NP →flying∩planes

以便解释“they are flying -- a plane”（NP- Verb-NP）、“（flying）planes – are—noisy”（NP—Verb—Adjective）这样的句子。不过，这个规则集合给我们提供句子“they are flying planes”的两个非等同推导式，归结为如下图表：

（24）#∩Sentence∩#



#∩Sentence∩#



因此，这个句子有两个被指派的短语结构：既可以分析为“they—are—flying planes”，也可以分析为“they—are flying—planes”。实际上，这个句子也正是以这种方式表现出歧义的。

3.4 英语语法的这些部分在很多方面被过分简单化了。一方面，（20）的每个规则和（23）的左端只有一个单一符号，尽管我们在§3.2 并没有对[∑, F]语法施以同样的限制。形如（25）的一个规则表明：X 只有在语境 Z—W 下才可以被改写成 Y。

（25） Z∩X∩W →Z∩Y∩W

很容易看出，如果我们允许这样的规则，英语语法将大大简化。在§3.2 中我们要求在形如（25）这样的规则中，X 必须是一个单一的符号。这就保证了一个短语结构图可以从任何推导式中构建出来。假如我们定下规则并要求这些规则依次运用（在运用过序列的最后一个规则之后又开始运用第一个规则），并且如果我们区分必有规则（obligatory rule，当我们在规则序列中遇到时必须运用）与可选规则（optional rule，可用可不用），那么，英语语法也可以大大简化。这些修正并不能改变语法的生成能力，但它们能够带来相当程度的简化。

因重要性起见做出某种承诺，要求语法实际上在有限的时间里生成大量的句子，看来似乎是合理的。更为确切地说，不可能徒然浏览规则序列（不运用规则），除非构筑的推导式的最后一行是终极符号串。我们可以通过对规则序列中的强制性规则加上某种条件来满足这种要求。我们将一部专有语法（proper grammar）定义为一个系统[∑, Q]，其中∑是初始符号串集合，Q 是一个如（18）中 $X_i \rightarrow Y_j$ 那样的规则序列，附加条件是：对于每个 i，必须至少有一个 j，使得 $X_i=Y_j$ 并且 $X_i \rightarrow Y_j$ 是一个强制性规则。这样，（18）中规则的每个左边的项必须至少出现在一个强制性规则中。这就是最弱的简单条件，保证每当我们浏览规则时，必须有一个非终端的推导式至少先行一步。该条件规定：如果 X_i 可以改写成 $Y_{i1},\cdots, Y_{ik}$ 之一，那么这些改写式中至少有一个必须出现。然而，专有语法从本质上讲不同于[∑, F]语法。令

D（G）为可以从短语结构语法 G 产生出来的推导式的集合，不论 G 合适与否。令 D_F={D（G）|G 为[∑, F]语法}，D_Q={D（G）|G 为专有语法}，那么，

（26）D_F 与 D_Q 不是可比的，即 $D_F \not\subset D_Q$ 且 $D_Q \not\subset D_F$。

这就是说，有些短语结构系统可以用[∑, F]语法描写但不可以用专有语法描写，有些短语结构可以用专有语法描写却不能用[∑, F]语法描写。

3.5 我们已经定义了三种语言：有限状态语言（见§2.1），可推导语言和终极语言（见§3.2）。这几种语言按照下列方式相关联：

（27）（i）每一种有限状态语言都是终极语言，但不能反言之；

（ii）每一种可推导语言都是终极语言，但不能反言之；

（iii）存在可推导的非有限状态语言和有限状态的不可推导的语言。

假定 L_G 是符合§2.1 中的有限状态语法 G 的有限状态语言。我们按照下列方式来构筑[∑, F]语法：∑={S_O} ; F 包含一个形如（28i）的规则，对于每个 i, j, k 而言，（S_i, S_j）∈G, j≠o, 且 k≤N_{ij}; F 包含一个形如（28ii）规则，对于每个 i, k 而言，（S_i, S_o）∈G，且 k≤N_{io}。

（28）（i）$S_i \rightarrow a_{ijk} \cap S_j$

（ii）$S_i \rightarrow a_{iok}$

明显地，源自这种[∑, F]语法的终极语言正是 L_G，证实了（27i）的第一部分。

在§2.2 中，我们看到（12）中的 L_1、L_2、和 L_3 不是有限状态语言，可是 L_1 和 L_2 是终极语言。例如对于 L_1，我们有[∑, F]语法：

（29）∑：Z

F：Z→a∩b

Z→a∩Z∩b

这就证实了（27i）。

假定 L_4 是可推导语言，具有词汇 V_p={a_1,…,a_n}。假定我们在 L_4 的语法加上规则 $a_i \rightarrow b_i$ 的有限集合，其中 b_i's 不在 V_p 中且完全不同。那么，这种新的语法给出一种终极语言，该终极语言只是 L_4 的一个符号变体。这样，每个可推导语言也是终极语言。

作为终极的不可推导语言的一个例子，请考虑仅含有符号串的语言 L_5：

（30）a∩b, c∩a∩b∩d, c∩c∩a∩b∩d∩d, c∩c∩c∩a∩b∩d∩d∩d,…

一个无限可推导语言必须含有一个符号串的无限集合，这些符号串必须按照如下方式排列成序列 S_1, S_2, … 即：对于某种规则 X→Y 而言，对每个 i>1 来说，S_i 都可以通过运用这个规则从 S_{i-1} 得出。而且在这个规则中，Y 必须通过用一个符号串替换 X 的<u>单一符号</u>（下划线为原文所加，下同）从 X 形成（比较（18））。这对 L_5 来说明显是不可能的。可是，这种语言按照下列给定的语法是终极语言：

（31）∑：Z

F：Z→a∩b

Z→c∩Z∩d

有限状态不可推导语言的一个例子是 L_6，L_6 包含所有由 a 的 2n 或 3n 次出现构成的符号串，并且仅包含这些符号串（n=1, 2,…）。（12）中的语言 L_1 是一个可推导的非有限状态语言，带有初始符号串 a∩b 和规则 a∩b→a∩ a ∩ b ∩b。

法则（27）的主要意义在于按照短语结构进行的描写本质上比按照自左至右产生句子的有限状态语法所进行的描写有力（不仅仅是简单）。在§2.3 中我们看到，英语完全超出了这些语法的界限，因为它跟（12）中的 L_1 与 L_2 共享镜像属性。可是，我们刚才看到，L_1 是终极语言而且 L_2 也是终极语言。因此，我们会考虑拒绝有限状态模型，但我们不会基于类似的考虑拒绝更为有力的短语结构模型。

我们注意到后者（短语结构语法）在下列意义上比有限状态模型更为抽象，即：未包含

在语言词汇中的符号进入了这种语言的描写中。根据§3.2，V_p 合理地包含 V_T。这样，就（29）而言，我们按照不在 L_1 中的成分 Z 来描写 L_1；就（20）－（24）而言，我们在描写英语结构中引入诸如句子、名词短语、动词短语等不是英语单词的符号。

3.6 我们可以按照下列方式将形如（18）的[∑, F]语法解释成一个相当初始的有限状态过程。考虑具有有限状态 S_o,…，S_q 的系统，当处在状态 S_o 时，系统能够产生∑的任何符号串，由此进入新的状态。系统在任何点上的状态决定于成分 $X_1 \cdots X_m$ 的子集，$X_1 \cdots X_m$ 作为子串包含于最后产生的符号串中；系统通过将规则之一运用于这个符号串上产生新的符号串而进入新的状态。系统返回状态 S_o 产生终端一个符号串。于是在§3.2 的意义上，这个系统产生了推导式。这一过程在任何一点上都决定于它目前的状态和最后产生的符号串；在这一过程能够继续进行产生在有限方面中的某一方面不同于其最后输出的新符号串之前，对这一符号串的检查次数施加有限上界限制是必要的。

构筑超出[∑, F]语法描写范围的语言并不困难。实际上，（12iii）中的语言 L_3 明显不是终极语言。我不知道英语实际上是否是终极语言，也不知道是否实际存在其他语言完全超出短语结构描写的范围。因此，我看不出取消基于对（3）的考虑提出的这种语言结构理论的方法。然而，当我们回到描写的复杂性问题时（比较（4）），我们发现有充足的理由得出结论：这种语言结构的理论从根本上讲是不充分的。我们现在研究一些这样的问题，即当我们试图将（20）扩展到整个英语语法时产生的问题。

4 短语结构语法之不足

4.1 在（20）中，我们只考虑了扩展（develop）动词成分即“took”的一种方式。但是，即使就固定的动词词干而言，仍有大量其他形式可以出现在上下文“the man — the book”中，例如“takes”、“has taken”、“has been taking”、“is taking”、“has been taken”、“will be taking”，等等。直接描写这个成分集合是相当复杂的，因为这些成分之间的过分依赖（如“has taken”而非“has taking”，“is being taken”而非“is being taking”等）。实际上，我们可以给出很简单的分析，将“动词”分析成一个由相互独立成分构成的序列，只选择某些不连续符号串作为成分。例如，在短语“has been taking”中，我们可以离析出不连续成分“has …en”、“be…”和“take”；于是，我们可以说这些成分自由结合。系统地照此分析，我们用（32）替代（20）中的最后一个规则：

（32） （i）Verb→Auxiliary∩V
（ii）V →take, eat,…
（iii）Auxiliary→C（M）（have∩en）（be∩ing）（be∩en）
（iv）M →will, can, shall, may, must
（v） C→past, present

（32iii）中的符号可以做如下解释：在扩展推导式中的“助动词”（auxiliary）时，我们必须选择未加括号的成分 C，并且我们可以选择零或按照一定顺序选择多个加括号的成分。这样，在下面第五行连续进行（21）中 D_1 的推导过程中，我们可以如下进行：

（33）#∩the∩man∩verb∩the∩book∩# [见（21）D_1]
#∩the∩man∩Auxiliary∩v∩the∩book∩# [（32i）]
#∩the∩man∩Auxiliary∩take∩the∩book∩#［（32ii）]
#∩the∩man∩C∩have∩en∩be∩ing∩take∩the∩book∩#
[（32iii），选择成分 C, have∩en 和 be∩ing]
#∩the∩man∩past∩have∩en∩be∩ing∩take∩the∩book∩#［（32v）]

假定我们将 Af 类定义为包含语缀“en”、“ing”和成分 C；将 V 类定义为包含所有 V、M、“have”和“be”。于是，我们可以按照如下规则将（33）的最后一行转换成一个排序合适的语素序列（sequence of morpheme）：

（34）Af∩v →v∩Af∩#

将这条规则运用于（33）最后一行的三个 Af∩v 序列的每个之上，我们推导出：

（35）#∩the∩man∩have∩past∩#∩be∩en∩#∩take∩ing∩#∩the∩book∩#.

在§2.2 的第一段中，我们提到，一部语法包含一个规则集合（所谓形态音位规则），这些规则将语素符号串转换成音位符号串。英语的形态音位学有如下规则（我们使用常规的而非音位的正字法）：

（36）have∩past→had
be∩en→been
take∩ing→taking
will∩past→would
can∩past→could
M∩present→M
walk∩past→walked
take∩past→took

将这种形态音位规则运用到（35），可以推导出下列句子：

（37）the man had been taking the book.

类似地，规则（32）、（34）将给出陈述句中动词的所有其他形式并且只给出这些形式，有一个主要的例外下面要谈到（有几个次要的例外这里忽略不谈）。

然而，这个非常简单的分析在好几个方面超出了[∑, F]语法的范围。规则（34）尽管非常简单，却不能并合到[∑, F]语法里，因为其中没有不连续成分的位置。而且，为了将规则（34）运用到（33）的最后一行，我们必须知道“take”是一个 V 因而是一个 v。换句话说，为了运用这条规则，有必要不仅仅检查这条规则适用的符号串；有必要知道这个符号串的某些成分结构，或者同样地，检查这个符号串之前的某些行。既然（34）要求知道符号串的“推导史”，它就违背了§3.6 中所讨论的基本属性要求。

4.2 这种将动词短语简单分析成独立选择单位构成的序列超出了[∑, F]语法的范围，这一事实表明，[∑, F]语法非常有限，不能给出语言结构的真实面貌。对动词短语的深入研究进一步支持这一结论。对（32）在介绍的成分的独立性有一个主要限制。如果我们选择一个不及物动词（如“come”、“occur”等）作为（32）中的 V，我们不可能选择 be∩en 作为助动词。不可能有“John has been come”、“John is occurred”以及类似的短语。而且，成分 be∩en 不可能独立于短语“verb”这样的上下文被选择。在“the man—the food”这样的上下文中如果选择成分“verb”，那么我们在运用（32）时就受到不能选择 be∩en 的限制，尽管我们能自由地选择（32）的任何其他成分。就是说，可以有“the man is eating the food”、“the man would have been eating the food”等这样的短语，但不能有“the man is eaten the food”、“the man would have been eaten the food ”等这样的短语。另一方面，如果短语“verb”的上下文比方说是“the food—by the man”，那么就要求选择 be∩en。可以有“the food is eaten by the man”而不能有“the food is eating by the man”等。总之，我们看到：成分 be∩en 进入详细的限制网络使其跟（32）中分析“verb”时所介绍的所有其他成分区别开来。be∩en 的这种复杂独特行为表明：将 be∩en 从（32）中排除出去并以某种其他方式将被动式引入[∑, F]语法是可取的。

事实上，有一种很简单的方法将带有 be∩en（即被动式）的句子合并到[∑, F]语法里。我们注意到，对于每一个“the man ate the food”这样的主动句，都有一个相应的被动句“the food was eaten by the man”，反之亦然。设定我们将成分 be∩en 从（32iii）中删除，然后将下列规则加到[∑, F]语法上：

（38）如果 S 是一个形如 NP_1—Auxiliary—V—NP_2 的句子，那么该形式对应的

符号串 NP_2—Auxiliary∩be∩en—V—by∩NP_1 也是一个句子。

例如，如果“the man—past—eat—the food”（NP_1—Auxiliary—V—NP_2）是一个句子，那么“the food—past be en—eat—by the man”（NP_2—Auxiliary∩be∩en—V—by∩NP_1）也是一个句子。规则（34）与（36）将第一个符号串转换为“the man ate the food”，将第二个符号串转换为“the food was eaten by the man”。

这种关于被动式的分析的好处是不会出错。既然成分 be∩en 被从（32）中删除，那么使（32）适合上文所讨论的复杂限制就不再必要。下面这些事实现在看来，在每一种情况下都是我们刚才所给分析的自动结果：be∩en 只能和及物动词一起出现；在“the man—the food”这样的上下文中被排除；在“the food—by the man”这样的上下文中要求出现。

可是，形如（38）的规则刚好超出了短语结构语法的限制。像（34），将它所适用的符号串的成分重新排序，并且它要求相当数量的关于这个符号串的成分结构的信息。当我们对英语句法作进一步详细研究时，我们发现，在很多其他情形下，如果[∑, F]系统得到跟（38）一样的普遍形式规则的补充，则这种语法可以简化。我们不妨称每个这样的规则为一个语法转换式（grammatical transformation）。作为描写语言结构的第三种模型，我们现在来简单考虑转换语法的形式属性，转换语法可以跟短语结构的[∑, F]语法邻接起来[⑧]。

5 转换语法

5.1 每个转换式 T 从本质上讲都是一条规则，将每个带有给定成分结构的句子转化成一个新的带有推导成分结构的句子。对于每个转化 T 来讲，转换式与其导出结构必须以一种固定不变的方式跟被转换的符号串结构相联。我们可以用结构的术语将 T 的属性描述为它所运用的符号串的域（domain）和它在任何那种符号串上所实现的变化。

在下面的讨论中我们像§3.2 那样假定：有一个[∑, F]语法带有词汇 V_P 和一个终端词汇 $V_T \subset V_P$。

在§3.3 中我们指出，[∑, F]语法允许终极符号串的推导；并且我们指出，一般说来，一个给定的符号串往往有好几个等同的推导式；两个推导式如果能规约成相同的形如（22）那样的树形图，则我们所它们是等同的，等等[⑨]。设定 $D_1 \cdots D_n$ 一个终极符号串 S 等同推导式的最大集合。于是我们可以将 S 的短语标记（phrase marker）定义为作为推导式 $D_1 \cdots D_n$ 之列出现的符号串的集合。一个符号串具有不止一个短语标记，当且仅当：该符号串具有非等同推导式（比较（24））。

设定 K 是 S 的一个短语标记，我们说：

（39）（S,K）可以分析为（$X_1, \cdots, X_n$）当且仅当：存在符号串 $s_1, \cdots, s_n$ 使得：

（i）$S = s_1 \cap \cdots \cap s_n$

（ii）对于每个 $i \leqslant n$，K 包含符号串 $s_1 \cap \cdots \cap s_{i-1} \cap X_i \cap s_{i+1} \cap \cdots \cap s_n$

（40）在这种情形下，关于 K，s_i 是 S 中的一个 X_i[⑩]

（40）中定义的关系刚好是§3.3 中定义的“是一个”（is a）关系；也就是说，s_i 在（40）意义上是一个 X_i 当且仅当：s_i 是 S 的一个子串，S 可以追溯到形如（22）树形图的一个单一节点，这个节点标记为 X_i。

上面定义的“可分析性”（analyzability）概念容许我们准确的具体指明运用任何转换的范围。跟每一个转换相联，我们给出一个限制类 R，定义如下：

（41）R 是一个限制类（restricting class）当且仅当：对于某些 r, m，R 是下列序列的集合：

---

⑧ 关于语言描写转换式的代数学的详细发展及转换语法的报道，请看参考文献[3]。有关将这类描写进一步运用到语言材料上，请看[1]、[2]，以及角度有某种不同的[4]。

⑨ 尽管相当繁冗，但却不难给出所讨论的等同关系的严格定义。

⑩ “是一个”（“is a”）的概念实际上应该进一步关系化为 S 中 s_i 的一次给定出现。我们可以将 S 中 s_i 的一次出现定义为一个有序对（s_i，X），其中 X 是 S 的初始子串，s_i 是 X 的最后子串，参看[5], P297。

$X_1^1, \cdots, X_r^1$

⋮

$X_1^m, \cdots, X_r^m$

其中，对于每个 i, j 而言，X_i^j 是词汇 V_P 的符号串。这样，我们说，如果跟转换 T 相联的限制类 R 包含（S, K）可以分析成的序列（X^j_1, ⋯, X^j_r），则带有短语标记 K 的符号串 S 属于转换 T 的范围。这样，转换的范围是符号串 S 的有序对（S, K）和 S 的短语标记 K 的集合。就一个带有歧义成分结构的符号串 S 而言，转换可以运用于带有一个短语标记的 S，但不运用于带有第二个短语标记的 S。

特别地，（38）描写的被动转换使其跟仅含有一个序列的限制类 R_P 相联：

（42）R_P={（NP, Auxiliary, V, NP）}

这个转换可以适用于任何这样的符号串：一个名词短语后面跟着一个助动词，助动词后面跟着一个动词，动词后面跟着一个名词短语。例如，这个转换可以运用于符号串（43），（43）按照短杠可以分析成子串 $S_1, \cdots, S_4$：

（43）the man—past—eat—the food

5.2 按照这种方式，我们可以用结构术语描写运用了任何转换的符号串（带有短语标记）的集合。现在，我们必须具体说明转换在其范围内对任何符号串实现的结构变化。一个初始转换 t 可以用下列属性加以界定：

（44）对于每个整数对儿 n, r（n≤r），存在唯一的整数序列（$a_0, a_1, \cdots, a_k$）和 V_P 上唯一的符号串序列（$Z_1, \cdots, Z_{k+1}$）使得：

（i）$a_0=0$; k≥0; 对于 1≤j≤k 来说，1≤a_j≤r；$Y_0=U$[11]

（ii）对于每个 Y_1, ⋯, Y_r t（$Y_1, \cdots, Y_n$; Y_n, ⋯, Y_r）= $Y_{a0}\cap Z_1\cap Y_{a1}Z_2\cap Y_{a2}\cap\cdots\cap Y_{ak}\cap Z_{k+1}$

于是，t 可以理解成将上下文（45）中 Y_n 的出现转化成一定的符号串 $Y_{a0}\cap Z_1\cap\cdots\cap Y_{ak}\cap Z_{k+1}$

（45）$Y_1\cap\cdots\cap Y_{n-1}\cap\text{--}\cap Y_{n+1}\cap\cdots\cap Y_r$

若给定 $Y_1\cap\cdots Y_r$ 分解成的项目序列（Y_1, ⋯, Y_r），则 $Y_{a0}\cap Z_1\cap\cdots\cap Y_{ak}\cap Z_{k+1}$ 是唯一的。t 将符号串 $Y_1\text{-}\cdots\text{-}Y_r$ 转化成以固定的方式跟它相联的新的符号串 $W_1\text{-}\cdots\text{-}W_r$。更确切的说，我们将导出转换 t^* 跟 t 联系起来：

（46）t^* 是 t 的导出转换，当且仅当：对于所有的 $Y_1\text{-}\cdots\text{-}Y_r$，t*（$Y_1\text{-}\cdots\text{-}Y_r$）=$W_1\cap\cdots\cap W_r$，其中，对于每个 n≤r 来说，$W_n$= t（$Y_1, \cdots, Y_n$; $Y_n, \cdots, Y_r$）

现在，我们将初始转换 t 跟每个转换 T 联系起来。比方说，将初始转换 t_P 跟被动转换（38）联系起来，定义如下：

（47）t_P（Y_1; $Y_1, \cdots, Y_4$）=Y_4

t_P（Y_1, Y_2; Y_2, Y_3, Y_4）=$Y_2\cap$be$\cap$en

t_P（Y_1, Y_2, Y_3; Y_3, Y_4）=Y_3

t_P（$Y_1, \cdots, Y_4$; Y_4）=by$\cap Y_1$

t_P（$Y_1, \cdots, Y_n$; $Y_n, \cdots, Y_r$）=Y_n （n≤r≠4）

导出转换 t*于是具有如下效力：

（48）（i）t*（$Y_1, \cdots, Y_4$）=$Y_1 - Y_2\cap$be$\cap$en – Y_3 - by$\cap Y_1$

（ii）t*（the ∩man past eat the ∩food）= the ∩food – past ∩be∩en – eat - by∩the ∩man

规则（34）、（36）将（48ii）的右端变成“the food was eaten by the man”，正像它们将（43）变成相应的主动句“the man ate the food”一样。

[11] 其中 U 是一个同样成分，参注释 4。

如（38）所描写的那样，（42）、（47）所示的序偶（R_P, t_P）完全刻画了被动转换的特征。R_P 告诉我们该转换施用于哪些符号串（给定这些符号串的短语标记）和如何细分这些符号串以便运用这种转换，t_P 告诉我们在被细分的符号串上实现的结构变化。

一个语法转换由一个限制类 R 和一个初始转换 t 得以完全的具体说明，R 和 t 都可以像被动式的情况那样进行有限特征描绘。像上文概略的那样严格定义这种说明的方式并不困难。为了使转换语法的发展完整，有必要说明一个转换是如何自动将一个导出短语标记指派给每个转换式，有必要针对有关符号串集合的转换进行概括归纳。（这些内容和相关的论题在文献[3]中加以处理。）这样看来，一个转换就是将一个带有短语标记 K 的符号串 S（或者一个由那样的序偶构成的集合）转化成一个带有导出短语标记 K’的符号串 S’。

5.3 基于上述考虑。我们将语法描绘成一个三元结构。跟短语结构分析对应，我们有一个形如 X→Y 的规则序列，如（20）、（23）、（32）。从这一点出发，我们有一个形如（34）和（38）这样的转换规则序列。最后，我们有（36）那样的形态音位规则的序列，（36）又是具有 X→Y 这样的形式。为了从这样的语法中生成一个句子，我们构建一个扩充的推导式，以短语结构语法的一个初始符号串比方说（20）中的#∩Sentence#开始。然后，我们浏览短语结构规则，生成终极符号串。然后运用一定的转换，以正确的顺序给出一个语素符号串，这个语素符号串也许跟原来的终极符号串大不相同。运用形态音位规则将这个符号串转化成一个音位符号串。我们可以多次浏览短语结构语法，然后对结果的终极符号串的集合运用综合转换。

在§3.4 中，我们注意到，将短语规则排成序列和区分强制性规则和选择性规则是有好处的。同样，设立语法的转换部分也是有好处的。在§4 中，我们讨论了转换（34）和被动转换（38）。（34）将序列“词缀－动词”（affix—verb）转化成“动词－词缀”（verb—affix）。注意，（34）在每个扩充推导式中都必须运用；否则，将产生不合法的句子。这样，规则（34）是一条强制性转换。然而，被动转换可用可不用，两种方式都可得到一个句子。因此，被动转换是一条选择性转换。选择性转换与强制性转换之间的这种区别使我们在语言的两类句子之间做出区分。一方面，只运用强制性转换从短语结构语法的终极符号串推导出基础句子的核心（kernel）；然后，在作为核心句子基础的符号串上运用选择性转换生成导出句的集合。

当实际对英语结构做详细研究时，我们发现，如果将核心句限制在一个非常小的简单、主动、陈述句如“the man ate the food”等的集合范围内（实际可能是有限集合），语法可以大大简化。这样，我们可以通过转换推导出问句、被动句、带联结词的句子、带复合名词短语的句子（如“proving that theorem was difficult”带有名词短语“proving that theorem”）[12]。既然转换的结果是一个带推导成分结构的句子，则转换可以是复合的，我们可以从被动句推导出问句（如“was the food eaten by the man”），等等。现实生活中的实际句子通常并不是核心句，而是这些核心句的复杂转换式。可是，我们发现，转换大体上是“保留意义的”（meaning—preserving），以便我们能够将核心句看作某个给定句子的基础，这个给定的句子在某种意义上是初始的“内容成分”，现实的转换据此被“理解”。我们将在§6 简单谈论这一问题，更为广泛的讨论见文献[1]、[2]。

在§3.6 中我们指出，短语结构语法是有限状态过程的一种相当初级的类型，有限状态过程决定于其当前每一点的状态及其有限的最后输出。在§4 中我们看到：这种限制是非常严格的；通过加上考虑一定数量成分结构（也就是一定的推导史）的转换规则，语法可以被简化。然而，每种转换仍可进行有限特征描绘（参§§5.1—2），跟转换相联的有限限制类（41）显示出为了运用这种转换需要多少有关符号串的信息。因此，语法仍可被看成跟短语结构相

[12] 注意，这个句子要求一个综合转换作用于带有各自短语标记的一对符号串上。这样，我们有一个转换将形式分别为 NP—VP_1, it—VP_2 的 S_1、 S_2 转化成符号串 ing∩VP_1 – VP_2。这个转换将 S_1=“they – prove that theorem”, S_2=“it – was difficult”转化成“ing prove that theorem – was difficult,”，通过（34）进而变成“proving that theorem was difficult”。详见文献[1]、[3]。

对应的初级有限状态过程。即使不仅仅必须知道最后的输出（推导的最后一行），但对于每个语法来说，为了使推导过程继续下去，必须查阅多少过去的输出，仍然有一个界限。

6 语言理论的解释力

到现在为止，我们只根据简明性（simplicity）这种本质上的形式标准来考虑语言结构理论的相对充分性。在§1 中我们暗示，存在其他对于这种理论充分性的相关考虑。我们可以问（参（5））：这些理论所揭示的句法结构是否提供了对于语言使用（use）与语言理解（understanding）的研究。这里，我们不可能探讨这一问题。不过，即使是简单的谈论也将表明：这一标准为我们已经考虑过的三种模型提供了相同的相对充分性顺序。

如果一个语言的语法要提供对于语言理解方式的研究，则必须特别强调的是：若一个句子是歧义的（按照不止一种方式理解），那么这种语法要给这个句子提供两者选一的分析。换句话说，如果某个句子 S 是歧义的，我们可以这样来检验已知语言理论的充分性，即看根据这种理论为所研究语言构建的最简语法是否自动提供生成句子 S 的不同方式。根据这种检验将马尔可夫过程、短语结构和转换模型加以比较是有启发性的。

在§3.3 中我们指出，为英语设计的最简的[∑, F]语法刚好为实际上有歧义的句子“they are flying planes”提供了不同的推导式。可是，这一论断似乎并不适于有限状态语法。就是说，在任何可能被建议为英语一部分的有限状态语法中，没有明显的动因为这个歧义句指派不同的（推导）路径。这种同形异构（constructional homonymity）的例子（还有其他的例子）构成短语结构模型优于有限状态语法的独立证据。

对英语的进一步研究遇到一些琐碎的例子用短语结构模型不容易解释，请看短语（49）：

（49）the shooting of the hunters

我们可以将这个带“hunter”的短语理解成主语，类同于（50），或者将其理解成宾语，类同于（51）：

（50）the growling of lions

（51）the raising of flowers

不过，（50）、（51）并非类似地也是歧义的。当然，根据短语结构，这些短语的每一个都可以表示为：the – V ∩ing – of ∩NP.

对英语的仔细分析表明：假如将短语（49）－（51）从核心句中剔除出去并通过转换重新引入这些短语，即通过转换 T_1 将“lions growl”这样的句子转换成（50），通过转换 T_2 将“they raise flowers”这样的句子转换成（51），那么我们可以简化语法。当被正确构建的时候，T_1 和 T_2 跟注释 12 所描写的名物化转换相似。但是，“hunters shoot”与“they shoot the hunters”都是核心句；对前者运用 T_1，对后者运用 T_2 都产生结果（49）。因此，（49）有两个不同的转换起源。在转换层级上，（49）是一个同形异构的例子。（49）中语法关系的歧义性是下面这个事实的结果：在作为基础的核心句中“shoot”与“hunters”之间的关系不同。在（50）、（51）的情形下，就没有这种歧义性，因为无论是“they growl lions”还是“flowers raise”，都不是合乎语法的核心句。

还有其他很多同样具有一般性的例子（参[1]、[2]），在我看来，这些例子不仅为语言结构的转换观念的更大充分性，而且为§5.4 所表达的观点即转换分析可以使我们部分地将解释如何理解一个句子的问题简化为解释如何理解核心句的问题，提供了非常有力的证据。

总结：我们可以将一个语言描写成具有一个小的、可能有限的基础句的核心。基础句的核心带有§3 意义上的短语结构，与之相伴的还有一个转换的集合。这些转换可用于核心句或早先的转换式，从初始成分（elementary component）产生新的更为复杂的句子。我们已经看到某些迹象，说明这一方法可以使我们将实际语言的巨大复杂性简化为可以操作的程度；此外，这一方法可以对语言的实际应用与理解提供相当大的洞察力。

参考文献

[1] Chomsky, N.， The Logic Structure of Linguistic Theory（mimeographed）

[2] Chomsky, N., Syntactic Structure to be published by Mouton & Co., S-Gravenhage, Netherlands

[3] Chomsky, N., Transformational Analysis, ph. D. Dissertation, University of Pennsylvania, June, 1955

[4] Harris, Z. S.， Discourse Analysis Language 28, 1 （1952）

[5] Quine, W. V., Mathematical logic, revised edition, Harvard University Press, Cambridge, 1951

[6]Rosenbloom, P., Elements of Mathematical logic, Dover, New York 1950

[7] Shannon & Weaver, The Mathematical Theory of Communication, University of Illinois Press, Urbana, 1949

原文载 IRE Transaction on Information Theory, IT-2, pp. 113-124, Proceedings of the Symposium on Information Theory, Sept, 1956.