

《句法结构》

图书基本信息

书名：《句法结构》

13位ISBN编号：SH9190-004

10位ISBN编号：SH9190-004

出版时间：1979/7第1版1984/7第2次印刷

出版社：中国社会科学出版社

作者：(美)诺姆-乔姆斯基

页数：132

译者：黄长著,林书武,庞秉均,邢公畹

版权说明：本站所提供下载的PDF图书仅提供预览和简介以及在线试读，请支持正版图书。

更多资源请访问：www.tushu000.com

《句法结构》

内容概要

美国语言学家乔姆斯基 (Chomsky) 著，1957年荷兰海牙摩顿出版公司出版。全书共10章。在此之前，乔姆斯基于1955年写了《语言理论的逻辑结构》，由于种种原因，该书直到1975年才正式出版，《句法结构》可以说是该书的纲要。

《句法结构》是对美国传统的描写语言学的一场革命，其革命性首先表现在乔姆斯基站在理性主义的立场上反对描写语言学的经验主义。

他把语言学看成跟自然科学中的其他科学一样，可以从假设出发，进行推演并形式化。换句话说，非经验主义是可能的。《句法结构》有一半篇幅用于英语语法的形式化。非经验主义和形式化是转换生成语法的首要标志。

把句法关系作为语言结构的中心并以此说明语句的生成是这场革命的又一表现。为了描写和解释语言现象，乔姆斯基在《句法结构》中论证了语法的生成能力，认为应该把语法看成是能生成无限句子的有限规则系统。

它以"核心句"为基础，通过转换规则描写和分析不同句式之间的内在联系。该书分析了以"马尔可夫过程"为基础的通讯理论，认为它只能生成有限状态的语法，而这种"有限状态的语法"不能生成象英语这种语言里含有不连续结构的所有合乎语法的句子。基于此，乔姆斯基提出了转换语法模式，认为它才能生成所有合乎语法的句子而不会生成不合乎语法的句子。转换语法模式由短语结构规则、转换规则、语素音位规则三套规则构成。

短语结构规则有三种：合并、递归、推导式，其基本形式是 $x \rightarrow y$ 。读作"改写"，这个公式就是将x改写成y。短语结构规则生成的是"核心语符列"，不经过转换直接由这种语符列得出的基本句型叫"核心句"。

转换规则包括：移位、删略、添加。最后运用语素音位规则得出实际说出的句子。这三套规则中，最引人注目的是转换规则，因为短语结构规则和语素音位规则实际上继承了描写语言学的"直接成分分析"和语素音位的分析，转换是一种创新，它使语法具有更强的解释力。

《句法结构》把语义排除在语法之外，这一时期的理论框架不包括语义部分。乔姆斯基认为，语法理论不应该建立在语义的基础上，而应该用某种严格的、客观的方法去代替对于模糊的语义的依赖。不过这一立论在后来的发展中做了重大的修正。

《句法结构》是转换生成语法第一阶段--古典理论的代表作，它的出版被人们认为是"乔姆斯基革命"开始的标志，也标志着转换生成语法的诞生。

《句法结构》

精彩短评

- 1、有些东西不是很懂.....
- 2、没有想象中的难懂。
- 3、早应该读一下。。。
- 4、没看完，以后再看
- 5、语言哲学的重磅炸弹
- 6、据说这个英文版比较好读~
- 7、乔帮主的革命性之作。
- 8、原作比教材上说的晦涩很多、但也在运算上浅白很多、很多反例都没考虑到、但终究还是神作啊。
- 9、乔姆斯基大人啊~~
- 10、硬着头皮，却又看不明白
- 11、...只能算作读过一部分..在 Chomsky 下自然语言成为了一个形式语言~
- 12、感觉从这本入手会清晰一点
- 13、牛逼
- 14、爱死了.....
- 15、大概还有10遍可以读的样子
- 16、无法形容的牛掰！
- 17、具体方法大多理解无能，只能理解核心意图
- 18、用逻辑符号把英文语法重写了一遍，好像没有看到一般的语言结构，是不是还有更好的书呢？
- 19、非逼着我去复印？
- 20、天才想法
- 21、做自然语言处理的理论基础——基于语料库的句子结构判断、POS、最大熵和马尔科夫链等预测模型。
- 22、推荐还是看英文版的吧
- 23、不得不说老乔的书真的很难读懂，得慢慢钻研才行，话说这本书意义巨大，很薄，但是在那个时代发出了另类声音，语言学必读。
- 24、能。。。给我一个简单版的看咩【虽然已经很薄【其实主要原因是数学不好
- 25、@_@
- 26、据说看本书前要把脑子中所有句法知识通通清除掉，否则此书就是一nightmare.
- 27、终于拜读了乔姆斯基传说中的大作《句法结构》，在这本书中乔爷用转换 - 生成规则，将语素音位规则和基本核心句结合起来，构建了一套庞大的且论证量巨大的句法结构，这些在书目最后列出的语法规则就相当于生产句子的机器，所有的英语句子都要遵守这些语法。乔爷用归纳的方法将语法给归纳出来，然后又把他视为一种类似于”理念性“的结构，试图用演绎的方式推出所有句子。乔姆斯基可以说是第一代程序猿，之后所有的程序猿都要按照这一规则，简直就是C语言的初级版，这也就是英语，如果乔姆斯基分析汉语，比如说”不知有一句妈卖批当讲不当讲“，那么这套理论就是一死，毕竟不考虑内容而只顾形式的做法确实有点困难，也难怪本维尼斯特非得杠乔姆斯基了。
- 28、虽然是语言学方面的书，但是其研究方法可以适用于更广阔的领域，用结构主义来研究实证主义很难涉及的范畴，为今后的研究开辟出了新的方法论。
- 29、计算理论（有限状态机，上下文无关文法，图灵机）解读和形式化语言学，其实就是机器学习的开始和编译原理的基础。某种语言符合多项式根的结构。
- 30、乔桑，我恨你。
- 31、经过无数二手文献的洗礼后，可算是把这个翻译版给读完了
- 32、人家和我说这本是乔姆斯基写的最容易懂的一本，然后我就凌乱了。
- 33、不感兴趣呀。。。
- 34、书薄逻辑清晰是看着舒服的两大要件。。
- 35、难啃，经典~

《句法结构》

- 36、#形式逻辑之美#
- 37、什么是时态？时间是打在be动词上的一个符号吗？
- 38、看得我云里雾里的，但我看到这确实是个伟大的想法

《句法结构》

精彩书评

1、是一本内容是关于语言分析，方法是利用数学的公理化方法，用抽象（共相，投影）思维来理解语言问题，：语言是一个空间结构的东西，语法就是公理，句子就是算子或者是变换，句子然后给你分类，利用商空间来构造句子，能符合语法的是一个空间，然后在空间中利用语法，分析句子元素的序列。最重要的解释了：什么是逻辑真-----符合语法规则，什么是语义真-----符合语义；例子就是：诗人得人比喻句，语法真（形式真），但是不是意义真共相---抽象-----其实啊，这本书告诉你的是为什么柏拉图反对诗人的问题，因为诗人写诗是逻辑真，语义上是假。这个就是比喻，在诗歌里是必须的和允许的。而柏拉图希望从形式（逻辑）和内容合一的东西，里推导整个世界就是从共相投影给世界，世界是相的投影。。。。这个就是柏拉图的相论。有限句法如何生成无限符合语法的句子？想起来了那个从一到无穷大的世界最大的作家就是一台装有26个字母机器，加入封闭圈，就是无限了？然后利用马科夫链来生成有限自动机生成无限的东西。。。。

章节试读

1、《句法结构》的笔记-第5页

2.2

On what basis do we actually go about separating grammatical sequences from ungrammatical sequences? I shall not attempt to give a complete answer to this question here, but I would like to point out that several answers that immediately suggest themselves could not be correct. First, it is obvious that the set grammatical sentences cannot be identified with any particular corpus of utterances obtained by the linguist in his field work. Any grammar of a language will project the finite and somewhat accidental corpus of observed utterances to a set (presumably infinite) of grammatical utterances. In this respect, a grammar mirrors the behavior of the speaker who, on the basis of a finite and accidental experience with language, can produce and understand an indefinite number of new sentences. Indeed, any explication of the notion "grammatical in L" (i.e. any characterization of "grammatical in L" in terms of "observed utterances in L") can be thought of as offering an explanation for this fundamental aspect of linguistic behavior.

这里，乔氏提出了一个很重要的概念：有限语言vs无限语言。所谓有限语言是指 "any particular corpus of utterances"，亦即，收集来的语言素材、语料库等。对有限语言，确认句子的合法性是查看该句子是否在所收集的语料库中。但自然语言是无限语言，是无法用一个特定的话语语料库来囊括所有该语言的句子。

这里，我们需要区分几个术语：sequence、sentences，utterance。

sequence：是指有待判断合法性的字符序列（书面）或者一段发音（口头）。例如“寺佛人打尅”（字符序列），或者 "kamakura"（一段发音）。

sentence：通常是指通过合法性检验的字符序列。例如“人过大佛寺”就是一个sentence，而上述的“佛人寺打尅”就不是。sentence有时还指语料库中已经登录的、可以反复引证的句子。

utterance：是指“话语事件”。例如“人过大佛寺”，我重复了三遍，“人过大佛寺”，“人过大佛寺”，“人过大佛寺”。我只说了一个sentence，但我说了三个utterance。更重要的，utterance代表实际的语言运用活动，而sentence代表人的语言能力。对于讲母语的人来说，有些sentence我虽然没有听过、没有说过，但别人说了、写下来了，我既听得懂、自己也能说。而每一次的谈话，在网上发言就是我的utterance。

乔氏在这里第一次明确提出，说话者的语言行为是根据对该语言有限而偶然的经验反映的是语法，根据语法可以产生和理解无限多的新句子。（在后来的“标准理论”中，将behavior改成了competence）

。

2、《句法结构》的笔记-第8页

3 An Elementary Linguistic Theory

（图省略）

给定一篇文章，这篇文章由10个句子组成。现在的问题是，什么样的语法手段/装置，可以生成这些句子？例如一个拼音输入法软件，当你输入一串拼音时，这个软件怎样才能把这些拉丁字符串变成所期望的、合法的汉语句子或短语？如果是从拼音的拉丁字母直接转换，恐怕几乎是不可能的。因为这些串丢失了太多的构成汉语合法句子/短语所需要的信息，首先，没有四声信息，一个shi字符串对应了数十个汉字；第二，就算是有了四声选择范围仍然很宽。所以，从拼音转换成备选的汉字本身就富有挑

战性。从汉语的语言学层次来说，拼音层次和汉字层次是两个不同但相互关联的层次。有了汉字层次还是不够，还要有单词层次，词组层次，句子层次，甚至比句子更高的层次。那么，需要什么样的语法才能生成这些句子而不是非法的字符串？第一，分析这些层次的语法手段必须是有限的所以将这些句子收集起来当做是否符合语法的样板是行不通的，因为如果我通过增删改等手段修改了这篇文章，原来的的样板句就不起作用了。输入法软件也一样，它不可能事先预测用户将要输入什么，所以不可能将所有的中文句子都储存起来，因为和语法的中文句子数量是无限的。

有一种语言模型（有限状态机）可以生成一些简单的字符序列。这机器模型的工作原理是这样的，对于要输入的字串，他都有一个可对应的内部状态，比如输入法软件，你可能输入的字符就是a-z的小写字母。如果你输入其他字符，软件通常是拒绝转换。这样，输入法软件内部储存有26种合法的内部状态。除了这些合法的内部状态外，还有两个特殊字符，第一个是激活软件输入中文状态的键（mac机上是command+space键），第二个是当你选定了汉字字符串后用空格键。所以，我们把CMD/SP也称作是合法的内部状态，但和其他的合法内部状态不同，它们只能用在字符串的开头与结尾。这样一个线性产生字符串的装置我们称作有限状态语法，由有限状态语法产生的语言叫做有限状态语言。有限状态语法可以通过图表示。

3、《句法结构》的笔记-第7页

2.4 Third, the notion "grammatical in English" cannot be identified in any way with the notion "high order of statistical approximation to English." It is fair to assume that neither sentence (1) nor (2) (nor indeed any part of these sentences) has ever occurred in an English discourse. Hence, in any statistical model for grammaticality these sentences will be ruled out on identical grounds as equally 'remote' from English. yet (1), though nonsentential, is grammatical while (2) is not. Presented with these sentences, a speaker of English will read (1) with a normal sentence intonation, but he will read (2) with a falling intonation on each word; in fact with just the intonation pattern given to any sequence of unrelated words. He treats each word in (2) as a separate phrase. Similarly, he will be able to recall (1) much more easily than (2), to learn it much more quickly, etc. Yet he may never have heard or seen any pair of words from these sentences joined in actual discourse. To choose another example, in the context "I saw a fragile --", the words "whale" and "of" may have equal (i.e., zero) frequency in the past linguistic experience of a speaker who will immediately recognize that one of these substitutions, but not the other, gives a grammatical sentence. We cannot, of course, appeal to the fact that sentences such as (1) might be uttered some sufficiently far-fetched context, while (2) would never be, since the basis for this differentiation between (1) and (2) is precisely what we are interested in determining.

Evidently, one's ability to produce and recognize grammatical utterances is not based on the notion statistical approximation and the like. The custom of calling grammatical sentences those that 'can occur', or those that are 'possible', has been responsible for some confusion here. It is natural to understand 'possible' as meaning 'highly probable' and to assume that the linguist's sharp distinction between grammatical and ungrammatical is motivated by a feeling that since the 'reality' of language is too complex to be described completely, he must content himself with a schematized version replacing "zero probability, and all extremely low probabilities by impossible, and all higher probabilities by possible." We see, however, that this idea is quite incorrect, and that a structural analysis cannot be understood as a schematic summary developed by sharpening the blurred edges in the full statistical picture. If we rank the sequences of the given length in order of statistical approximation to English, we will find both grammatical and ungrammatical sequences scattered throughout the list; there appears to be no particular relation between approximation and grammaticality. Despite undeniable interest and importance semantic and statistical studies of language, they appear to have no direct relevance of determining or characterizing the set of grammatical utterances. I think we are forced to conclude that probabilistic models give no particular insight into some of basic problems of syntactic structure.

这一段的中心思想是：句子的合格性不能从统计方法得到。

在上一节，乔氏论证了句子的合法性不能从语义判断，为此乔氏还生造了一个没有意义的句子，重新

《句法结构》

抄录如下：

- (1) Colorless green ideas sleep furiously,
- (2) Furiously sleep ideas green colorless.

乔氏的观点是：尽管(1)这个句子在逻辑上矛盾，情理上荒唐，但并不影响它在语法上的合格性，因为这个(2)一比较就一目了然了。

在本段，基于同样的理由，同样不能用统计作为句子合格性的理由，因为以上述句子为例，无论什么样的语料库都不可能搜罗到像(1)那样的句子，也就是说(1)出现在语料库中的概率为零。乔氏的结论是“Evidently, one's ability to produce and recognize grammatical utterances is not based on the notion statistical approximation and the like.”

从生成语法的角度这是必然的，因为语法是一种装置，既能生成说话者已经说过的句子，也能生成从来没有说过的句子；既能生成有智慧、优雅的句子，也能生成逻辑荒唐、荒谬绝伦的句子。但是如果语言研究的目标是“语言使用”，亦即只研究人们已经说过、写过的句子，这个观点就值得商榷了。因为在现实生活中许多语言的使用是不合语法的。例如：

- (3) Fell in love of her but no idea how to chase her?

按照生成语法，明显是不合语法的，因为任何英语语法的生成装置是无法生成这样的句子的，识别装置已无法认识这样的句子。但如果语料库所收集的句子是在街头小广告或网络，这样的句子比比皆是。所以，乔氏认为生成语法只研究人的语言能力而不研究语言使用。

4、《句法结构》的笔记-第4页

注：我在第一页又加上了对该书前言的评注，请关注本笔记的朋友再次阅览一下第一页。Notice that in order to set aims of grammar significantly it is sufficient to assume a partial knowledge of sentences and non-sentences. That is, we may assume for this discussion that certain sequences of phonemes are definitely sentences and that certain other sequences are definitely non-sentences. This is a familiar feature of explication¹. 这里阐述的是生成语法的一般研究方法，演绎法：给定一定数量的“被观察的”对象，建立假说，并将其适用与该对象，看是否可以说明这些被观察对象。如果得到确认，则利用已建立的假说去预测新的事实或观察对象。如果仍然“适任”，再继续观察新的对象作出预测，如此反复。如果该假说不能正确反映新的事实或观察对象，则需修改假说。修改后的假说必须能够同时说明初始的观察对象和后来的新事实或观察对象。如果我们知道物理学的发展历史，不难看出，从牛顿力学到量子力学不也是循着这条轨迹发展过来的吗。尽管语言学比起物理学来远不成熟，一路走来跌跌撞撞，但我相信它正在沿着一条正确的轨道一点点前进。

以生成语法学派为代表的现代语言学50年来，经历了无数次的重大修改，甚至改的面目全非，层出不穷的新理论、新学说使人眼花缭乱，我觉得有两个重要的原因：第一：一精密科学为目标的现代语言科学仍然处在襁褓之中，人们虽然已经找到了明确的目标，但如何达到该目标并无明确的方法。各种理论都是假说，这些假说对纷纭复杂的语言现象还不能有一种完全系统的、普遍的、首尾一致的理论；第二：语言学的研究对象远比其他自然科学的研究对象复杂。因为语言是人类自身的产物，不同于完全外在于人类精神的客体。因此，对于语言学的研究对象甚至都没有搞清楚：乔氏理论自“标准理论”之后，认为语言学的研究对象应当是人类的“语言能力”(linguistic competence)，而不应当是“语言运用”(linguistic performance)。但是当代语言科学的应用、特别是以“自然语言处理”为代表的的应用，已经证明“语言运用”对于语言研究同样具有重大的意义。关于这一点，已经涉及语言哲学中的理性主义和实证主义的论战，无法在这里展开。但从“自然语言处理”领域出现的“基于规则的处理系统”和“基于语料库或基于统计的处理系统”正说明了我们对语言科学的研究对象并没有一个完全统一的、一致的结论。A certain number of clear cases, then, will provide us with a criterion of adequacy for any particular grammar. For a single language, taken in isolation, this provides only a weak test of adequacy, since many different grammars may handle the clear cases properly. This can be generalized to a very strong condition, however, if we insist that the clear cases be handled properly for each language by grammars all of which are constructed by the same method. That is, each grammar is related to the corpus of sentences in the language it

《句法结构》

describes in a way fixed in advance for all grammars by a given linguistic theory. We then have a very strong test of adequacy for a linguistic theory that attempts to give a general explanation for the notion of "grammatical sentences" in terms of "observed sentences" and for the set of grammars constructed in accordance with such a theory. It is furthermore a reasonable requirement, since we are interested not only in particular languages, but also in the general nature of Language. There is a great deal more that can be said about this crucial topic, but this would take us too far afield.

语言研究的目标是什么？就是在序言中所说，建立一套不参照个别语法的语言结构的理论。但这个建立过程，仍然要从个别语言的一批个别的“被观察的”句子开始。利用一些典型的例句是所建立的“假说式”语法能够适用于预测更多的新事实。但这只是万里长征的第一步。对于所观察的例句，有可能建立多个“适任”的语法。例如，对例句

(1) John hit the ball.

我们至少可以建立下述两套语法：

语法1：

- (a) S -> NP VP
- (b) NP -> Pron
- (c) VP -> V NP
- (d) NP -> Det N
- (e) Pron -> John
- (f) V -> hit
- (g) Det -> the
- (h) N -> ball

语法2：

- (a) S -> NP VP
- (b) VP -> V NP
- (c) NP -> Det N
- (d) NP -> Pron
- (e) Pron -> John
- (f) V -> hit
- (g) Det -> the
- (h) N -> ball

用这两套语法推导句子的生成过程也是不一样的。但最后生成结果却是一致的。到底哪一个语法更适任，从单个语言的测试也许很难找到答案。但如果能够对多个特定语言同时进新测试，而不改变语法本身，我们就可以得到一个更适任的语法，而且这个语法具有更广的普遍性。因为乔氏语言学理论的终极目标并不是为个别语言分别建立个别语法，而是为那个大写的"Language"，我的理解是所有特定语言的集合建立一套普遍的语言学理论，然后用这套语言学理论为每一个特定语言建立相应的语法。

5、《句法结构》的笔记-第2页

2. 语法的独立性

2.1

第一自然段：

中心议题：语言、句子、合语法性

什么是语言：语言是口语或书面语句子的集合（有限集合或无限集合）

词表(alphabet)：词的有限集合

《句法结构》

什么是句子：由有限个的词构成、通过语法生成的词串。

音素的数量是有限的，语素的数量是有限的，单词的数量是有限的，将单词连接起来构成比单词更大的语言对象的手段是有限的，但是由这些有限的元素所构建的句子的数量却是无限的。有限和无限的对立统一。为什么用有限的手段构建的语言对象却是无限的？

上述关于语言、语法、句子的定义与我们以前在学校学的“语言学知识”完全不同，不同在什么地方？这里所有的定义根据都是数学中集合的概念，是定量的，而我们以前学的关于上述概念的定义都是定性的，例如“语言是人类社会个体之间进行交流的工具”。乔氏是人类历史上第一次明确地用数学概念定义了以前被认为是文科的研究对象。

这些关于语言，语法，句子的定义，构成了以后的形式语言学和转换生成语法的理论基础，而形式语言学则是计算机科学中编译理论的理论基础。

在这里，乔氏提出：对语言进行分析的根本目的就是构筑语法，这套语法可以自动地将符合该语言规范（亦即合语法的）字串序列与不符合规范的（亦即不合语法的）的字串序列加以区分，并研究为什么给定的字串序列符合语法或者不符合语法。所以，在乔氏理论中，语法有两个目的：第一是生成手段（generative device），只生成和语法的句子而不生成不合语法的句子；第二是识别手段（recognition device），对给定的字串能自动辨别其合法性（grammaticality）。

6、《句法结构》的笔记-第3页

小结：

以上的几个自然段，乔氏重新定义了我们非常熟悉的一些语言学概念：

1. 什么是语言
2. 什么是语法
3. 什么是句子
4. 什么是句法

除此之外我们还遇到了乔氏引入的一些新的概念

1. 语言结构
2. 语言结构的理论
3. 语言学层次

关于这些概念的深入探讨不这片学习笔记的任务，我将在以后开栏专门探讨。

附注：国内许多人、特别是年轻人抱怨乔氏的著作晦涩难懂，读了半天不知所云。我想有两个方面的原因：第一，乔氏的英语行文经常是非常长的限定性从句加重读者阅读中的短期记忆负担，如果英语不过硬读了前面忘了后面，把整个句子串起来时对总体意思又把握不住；读中文也有类似的问题，冗长的限定性从句也使得译者不好处理，即不好分拆又不好不分拆，取其中，得到的译文就不伦不类，容易令人生厌；第二个原因也是最重要的，就像我开篇所说，乔氏理论中所使用的大部分术语仍然是我们日常常用的词汇，如grammar, language, sentence等，但和我们通常理解的已经完全不同。这种情况在其他学科也有。例如物理学中文的“做功”英文就是work。那这么一个最最普通的单词定义物理学如此重要的概念我想当年前被第一次翻译的时候也是要“晕到死”的吧。

由于《句法结构》是生成语法与计算机形式语言共同的经典，国内计算机界在翻译grammar时为了和我们日常用的“语法”相区别，使用了一个20，30年代流行但后来废止的词--“文法”以示区别。

7、《句法结构》的笔记-第1页

《句法结构》

《句法结构》读书笔记

《句法结构》在生成语法的发展历史中具有独一无二的地位，它是所有现代语言学理论包括形式语言学、转换生成语法的奠基石。读懂读通了《句法结构》就会对现代纷繁多样的各种语言学理论流派有了高屋建瓴、洞若观火的感觉。可以说《句法结构》的历史地位比乔氏后来的所有的语言学著作都要高。关于这个问题由于和学习《句法结构》关系不大在这里不展开谈，以后另辟专栏详谈。

学习方法：有条件的最好读英语原文，因为读乔氏的著作，读原文和读中文译文的难度大致差不多。

在学习《句法结构》之前，首先要“卸载”(unlearn)，亦即“忘掉”以前学到的所有语言学知识，包括：什么是语言，什么是语言学理论，什么是语法，什么是句法。不要再用对这些术语先入为主的观念去读这本书，否则将是一场噩梦。

0. 前言

这段内容是我新加上去的。原本对【前言】专门开个栏讨论，但时间不允许，只得放在这里了。对朋友们造成的不便深表歉意。

如果说《句法结构》这本书是不朽的名著的话，我觉得不在于它提出了转换规则（尽管在当时转换的概念被认为是“革命性的”），不在于它论证了有限状态语言（马尔科夫链）以及短语结构语法不适合描述像英语这样的自然语言，甚至不在于由于这本书开创了“转换生成语法学派”（因为总有一天会有新的语法理论代替它），而在于《句法结构》第一次，是人类历史上第一次提出了语言学理论必须形式化、必须向精密科学、自然科学演进的问题。这是后来50年多个新出现的学科包括计算机科学、理论语言学、计算语言学所有理论的正当化基础。乔氏的观点是，精确、严格的形式化的语言学理论，可以使人像在其他自然科学的理论中看到的那样，对所提出的理论经过实验检验，亦即，具有可证伪性。我摘抄一段我在网上看到的一段话：科学理论是这个样子的，它建立在有限的观察基础上，具有客观、逻辑、实证、可重复检验等特性，虽然不保证永远正确，但它是人类可以依赖的最可靠的知识。在科学理论的适用范围内，新出现的情况总是不断地验证它的正确性，它本身具有可以在将来被证伪的途径，一旦被证伪了，该理论就被推翻了，或者改变其适用条件成为2.0版。

这是语言学走向“语言科学”的历史性一步。现代语言学的定义是“语言学是对语言的科学性研究”（Lyons《Introduction to Theoretical Linguistics》1968）。“科学性（scientific）”原本的涵义就是“可重复试验、可证伪的过程”。乔氏认为“模糊不清、囿于直觉的概念既不能引出荒谬的结论，也不能提出新的、正确的结论”。

生成语法后来的发展背离了这个基本宗旨，从一门形式化科学走向了更为抽象的语法和人类心理映射关系的研究。而后来出现的GPSG和HPSG，则继续坚持这个发展方向，认为语言学最终应当向物理学看齐，从而可以称为一门具有可证伪性的科学理论。

1. 导论

第一自然段

中心议题：句法(syntax)、语法(grammar)、语言学理论

句法：句法是对句子构成的原则和过程的研究。

语法：句法研究的目的是建立语法，语法是产生所分析语言中所有句子的某种手段/装置。

在这里，句法是研究活动，而语法则这一研究活动的目的。

注意：在这里乔氏重新定义了语法的概念：乔氏将语法看作是生成该语言所有句子的一种“装置”，

《句法结构》

用现在的术语说，语法就是一套形式化的规则系统，这套系统就像数学公式一样，例如长方形面积公式 $A=lw$ 可以生成世界上所有长方形的面积，语言规则系统可以生成某一特定语言的所有句子。乔氏认为：语言学家所关注的问题应当是确定成功语法必要的最基本属性。也就是说，句法研究的目的就是回答这个问题：一套语法系统若要生成对象语言的所有句子，应当具备哪些特征、条件。

语言结构的理论：句法研究的最终结果应当是一套关于语言结构的理论，这套理论有两部分，第一部分：该理论可以在不参照目标语言的情况下抽象地表示或研究语法所使用的描述手段。例如句法规则的产生式 $S \rightarrow NP VP$ 就是一种描述手段，用这个产生式就可以部分地描述句子的构成原理；第二部分：在给定句库的情况下为所有语言选择语法提供一种通用的方法。产生式规则系统采用了重写形式 $A \rightarrow BC$ ，这就是一种通用的方法，一旦给定了特定语言的句库就可以为该语言提供具体的语法（注意这里语法的定义）。例如给定的语言是日语则句子的生成规则是 $S \rightarrow NP VP; VP \rightarrow NP V$ 。

第二自然段：

中心议题：语言学层次，语言学理论

语言学层次：句法只是语言学研究的一个部分：按照语言学研究的划分，应当分为语音学、音位学、形态学、句法学和语义学。为了在各个层次上研究语言，就产生了相应的结构，例如：音位结构、词结构，短语结构等。“语言层次本质上是一组用来构建语法的描述性手段”

人们说一段话，写一篇日记，就构成了一个discourse（语篇）单位。这个单位，可以分解为若干个句子（英语的句子有着明确的主谓关系，一个句子只能有一个带时态的谓语。汉语没有时态，古代汉语只有所谓的句读和断句；现代汉语受西方的影响，有了句子的概念，但句子和非句子之间的边界不是很清楚）。句子可以再分解为短语，短语再分解为单词（汉语中单词和非单词的界限也不是很清楚）。单词之下，可再分为语素，再往下分，就失去语义了，得到音素/音段和音位，成为语言分析的最小单位。在分析的每一个层级上，元素之间的结构都是不相同的，描述手段也不相同。这些在各个层次上的描述手段的总和，就是语法。这里，已经把语法的定义大大扩展了，几乎扩展到和语言学理论相同的规模。

语言学理论：语言学理论的任务就是要提供一套满足条件的语法系统，而这套系统应当包含上述的各个语言学层次。

“对语言结构这种纯形式化的研究对语义学的研究具有某种令人感兴趣的意义。”

所谓的“纯形式化”就是不考虑在某个层次上语言对象的结构背后是否代表其它的东西，例如，语义。例如，给定一个音位结构/blik/，分析它的结构，研究/l/的分布，而忽略是否真有某个发这个音的单词存在。

所以，生成语法所感兴趣的应当是语言对象的结构。

8、《句法结构》的笔记-第6页

2.3 Second, the notion "grammatical" cannot be identified with "meaningful" or "significant" in any semantic sense. Sentences (1) and (2) are equally nonsential, but any speaker of English will recognize that only the former is grammatical.

- (1) Colorless green ideas sleep furiously.
- (2) Furiously sleep ideas green colorless.

《句法结构》

Similarly, there is no semantic reason to prefer (3) to (5) or (4) to (6), but only (3) and (4) are grammatical sentences of English.

- (3) have you a book on modern music?
- (4) the book seems interesting.
- (5) read you a book on modern music?
- (6) the child seems sleeping.

Such examples suggest that any search for semantically based definition of "grammaticalness" will be futile. We shall see, in fact, in § 7, that there are deep structural reasons for distinguishing (3) and (4) from (5) and (6); but before we are able to find an explanation for such facts as these we shall have to carry the theory of syntactic structure a good deal beyond its familiar limits.

语义和语法的关系。按照生成语法的观点，语法决定语义。后者是派生的而前者是基本的、本质的。为什么？任何符号系统的外在表现形式都是“串”，不管是语音串还是字符串。而语义则是对这个“串”实例的解释（interpretation）。例如，对“手纸”这个字串，中国人的解读是“厕所用纸”，而日本人的解读则是“信件”。所以乔氏说“any search for semantically based definition of "grammaticalness" will be futile.”也就是说，用是否有意义来确定一个给定字串是否合法是行不通的。这是乔氏在这个段落想表达的中心思想。为此，乔氏不惜造出了一个“荒谬绝伦、不可理喻”的句子：

(1) Colorless green ideas sleep furiously. (无色的绿思想睡得疯狂)

这个句子后来成了一句经典，所有谈语言学的文章都在引用，和我们中国的“神马”、“浮云”、“木有”等有一比。

但乔氏在7年之后对这个观点进行了修正，认为语义在语法生成句子时确实“部分地”起一定的作用。这也就是那个著名的“深层结构的”由来。例如下面这两个句子：

- (1) 小王劝他的女朋友今晚不要离开他。
- (2) 小王答应他的女朋友今晚请她吃饭。

这两个句子的结构完全一样，都是主语+动词+从句。但(1)中从句的主语与主句的主语并不一致，而与主句的宾语一致；而(2)则正好相反：从句的主语与主句一致。这种现象显然不能通过语法结构上的排列组合来说明问题，而是和“劝”、“答应”词汇本身对句法的要求、亦即词汇的“句法意义、真值意义”在决定句子的句法结构。虽然乔氏后来承认了忽略语义的“错误”，但我看没有这个必要。因为乔氏在写《句法结构》的时候所设想的目标是语言描述的精密化、严格化和形式化，是要为语言建立数学模型。这时乔氏心目中的“语言”概念大概是最为宽泛的，除了自然语言外，也包括了像程序设计语言、逻辑语言和数学语言在内的人工语言。而在对这个广义上的“语言”进行形式化的研究时确实是不用考虑语义的。而对自然语言，如果只对其句法的形式侧面进行研究语义也不应是研究的对象。问题是乔氏后来把自己的研究定位成对“人类语言习得机制”、语言能力具有解释力的“普遍语法”(UG)，也就是要研究人类获取、掌握语言的心理机制，那么语言的语义侧面甚至语用侧面都要成为研究的对象。所以，是否将语义作为句法研究的一部分，应当是研究目标而定。关于乔氏认为例句中(5)和(6)不合法其实也不尽然。在非正式场合如私人小广告等经常看到像(5)那样的句子，如：

(3) Not sure where to find math tutor when you're stuck with your math assignment? here is where to turn for help.

既然生成语法是描述性语法(descriptive)而不是规定性语法(prescriptive)那就不应当排除像(3)这样的句子。

《句法结构》

版权说明

本站所提供下载的PDF图书仅提供预览和简介，请支持正版图书。

更多资源请访问：www.tushu000.com