Noam Chomsky and Generative Grammar

The simplest grammars discussed by Chomsky capable of generating an infinite set of stces by means of a finite number of recursive rules operating upon a finite vocabulary are Finite State Grammars (F.S.G.). Finite State Grammar proved less powerful for the description of syntax of natural lges. A lge comprises an infinite number of stces that can be extended indefinitely, but prachical limitations upon the lenght of any stce exist, though in principle no limit can be set to this length. eg: This is the man that married the girl that ... it can be expanded to the desired length. For Chomsky every different sequence is a different stce. In his grammar, every word must be assigned to a syntachic class (N=Nouns, V= Verbs ---etc).

1- Finite State Grammar

In Finite State Grammar we have different states: Initial State and Final State. In this gr, stces are generated by means of choices made from left to right according to a purely consecutive order moving from initial state to final state. Any sequence of words generated this way is grammatical. FSG generates a finite number of stces but it can be extended by adding loops. FSG cannot account for non-adjacent words e.g. The man / who says that / is lying, where the subject and the verb are not adjacent.

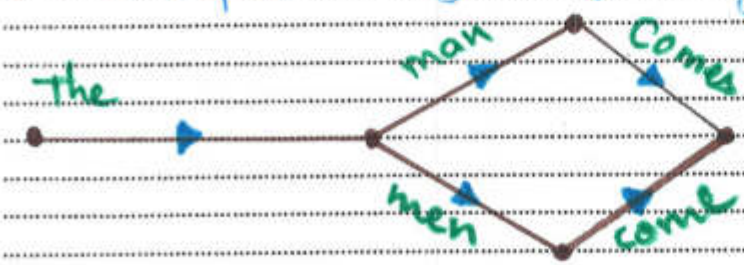
So embedded stces cannot be generated by FSG. In FSG all rules have been context free, every symbol is rewritten whenever it occurs.

Assuming the set of grammatical stces of English to be given, we now ask what sort of device can produce this set (similarly, what sort of theory gives an adequate account of the structure of this set of utterances). We can think of each stce of this set as a sequence of phonemes of finite length. A lge is an enormously involved system, and it is quite obvious that any attempt to present directly the set of grammatical phoneme sequences would lead to a gr so complex that it would be practically useless.

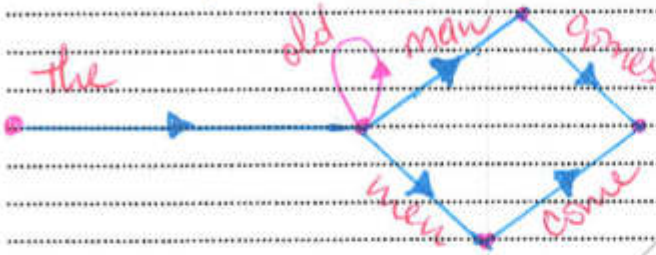
For this reason, linguistic description proceeds in terms of a system of "levels of representations". Instead of stating the phonemic structure of stces directly, the linguist sets up such "higher level" elements as morphemes and states separately the morphemic structure of stces and the phonemic structure of morphemes. It can easily be seen that the joint description of these two levels will be much simpler than a direct description of the phonemic structure of stces.

Let us now consider various ways of describing the morphemic structure of stces. We ask what sort of gr is necessary to generate all the sequences of morphemes (or words) that constitute grammatical English stces, and only these.

One requirement that a gr must certainly meet is that it be finite. Hence the gr cannot simply be a list of all morpheme (word) sequences, since there are infinitely many of these. Suppose that we have a machine that can be in any one of a finite number of different internal states, and suppose that this machine switches from one state to another by producing a certain symbol (let us say, an English word). One of these states is an intial state; another is a final state. Suppose that the machine begins in the initial state, runs through a sequence of states (producing a word with each transition), and ends in the final state. Then we call the sequence of words that has been produced a "sentence". Each machine thus defines a certain lge, namely the set of stces that can be produced in this way. Any lge that can be produced by a machine of this sort we call a finite state language, and we can call the machine itself a finite state Grammar. A FSG can be represented graphically in the form of a "state diagram". For example, the gr that produces just the two stces "the man comes" and "the men come" can be represented by the following state diagram:



We can extend this gr to produce an infinite number of stces by adding closed loops. Thus the finite gr of the subpart of English containing the above stces in addition to "the old man comes", "the old old man comes", ..., "the old men come", "the old old men come", ... can be represented by the following state diagram,



Given a state diagram, we produce a stce by tracing a path from the initial point on the left to the final point on the right, always proceeding in the direction of the arrows. Having reached a certain point in the diagram, we can proceed along any path leading from this point, whether or not this path has been traversed before in constructing the sentence in question. Each node in such a diagram corresponds to a state of the machine. We can allow transition from one state to another in several ways, and we can have any number of closed loops of any length.

This conception of lge is an extremely powerful and general one. If we can adopt it, we can view the speaker as being essentially a machine of the type considered. In producing a stce, the speaker begins in the initial state, produces the first word of the stce, thereby switching into a second state which limits the choice of the second word, etc.

Each state through which he passes represents the grammatical restrictions that limit the choice of the next word at the point in the utterance. For example, in the stce "the cat saw the dog" the gr would start by specifying the first word (i.e. selecting one of the set of possible first words for a stce in English, "The") then moves from this initial state to specify the next word (i.e. of the set of words that follow "The") and continues this process until the finite state has been reached.

Chomsky said that F.S.G is not a strong gr because:

1). It may lead us to generate an artificial lge, and linguistics is dealing with human or natural lge not artificial lge.

2). This extremely simple gr is incapable of accounting for many important processes of stce formation as in discontinuous constructions.

e.g: The boys who saw John are coming, where the grammatical relationship of the "the boys" to "are" cannot be handled in a F.S.G.