PRINCIPLED RESTRICTIONS ON THE
"THEORY OF TRANSFORMATIONS"

I. Some \( t \)'s FROM CHOMSKY (1957) AND THEIR DESCENDANTS IN LASNIK (1971).

1. \( \{ NP - C - V X \} \)
   \( \{ NP - C M - X \} \)
   \( \{ NP - C \} \) have \( - X \)
   \( \{ NP - C - b e - X \} \)
   \( X = X - X \), \( X = X \)
   \( X = X \), \( X = X \)
   \( X = X \), \( X = X \)
   \( X = X \), \( X = X \)

2. S.A. SAME AS 1.
   \( \{ x - X \}, \{ x - X \} \)
   \( \{ x - X \} \)
   \( \{ x - X \} \)
   \( \{ x - X \} \)
   \( \{ x - X \} \)
   \( \{ x - X \} \)
   \( \{ x - X \} \)
   \( \{ x - X \} \)

3. \( x - \{ S \} \)
   \( \{ M \} \)
   \( \{ \} \)
   \( \{ \} \)
   \( \{ \} \)
   \( \{ \} \)
   \( \{ \} \)
   \( \{ \} \)

4. \( \{ X - Y \} \) (WHERE \( X = Y \) OR \( Y = A \))
   \( \{ X - Y \} \)
   \( \{ X - Y \} \)
   \( \{ X - Y \} \)
   \( \{ X - Y \} \)
   \( \{ X - Y \} \)
   \( \{ X - Y \} \)
   \( \{ X - Y \} \)
   \( \{ X - Y \} \)

5. S.A. \( \{ X - Y \} \)
   \( \{ X - Y \} \)
   \( \{ X - Y \} \)
   \( \{ X - Y \} \)
   \( \{ X - Y \} \)
   \( \{ X - Y \} \)
   \( \{ X - Y \} \)
   \( \{ X - Y \} \)

6. \( \{ X - Y \} \)
   \( \{ X - Y \} \)
   \( \{ X - Y \} \)
   \( \{ X - Y \} \)
   \( \{ X - Y \} \)
   \( \{ X - Y \} \)
   \( \{ X - Y \} \)
   \( \{ X - Y \} \)

7. S.A. \( \{ X - Y \} \)
   \( \{ X - Y \} \)
   \( \{ X - Y \} \)
   \( \{ X - Y \} \)
   \( \{ X - Y \} \)
   \( \{ X - Y \} \)
   \( \{ X - Y \} \)
   \( \{ X - Y \} \)

8. S.A. \( \{ X - Y \} \)
   \( \{ X - Y \} \)
   \( \{ X - Y \} \)
   \( \{ X - Y \} \)
   \( \{ X - Y \} \)
   \( \{ X - Y \} \)
   \( \{ X - Y \} \)
   \( \{ X - Y \} \)

9. S.A. \( \{ X - Y \} \)
   \( \{ X - Y \} \)
   \( \{ X - Y \} \)
   \( \{ X - Y \} \)
   \( \{ X - Y \} \)
   \( \{ X - Y \} \)
   \( \{ X - Y \} \)
   \( \{ X - Y \} \)

B. ALL OF THE FOLLOWING ARE OPTIONAL AND UNORDERED:

10. \( \{ X - Y \} \)
    \( \{ X - Y \} \)
    \( \{ X - Y \} \)
    \( \{ X - Y \} \)
    \( \{ X - Y \} \)
    \( \{ X - Y \} \)
    \( \{ X - Y \} \)
    \( \{ X - Y \} \)

11. \( \{ X - Y \} \)
    \( \{ X - Y \} \)
    \( \{ X - Y \} \)
    \( \{ X - Y \} \)
    \( \{ X - Y \} \)
    \( \{ X - Y \} \)
    \( \{ X - Y \} \)
    \( \{ X - Y \} \)

12. \( \{ X - Y \} \)
    \( \{ X - Y \} \)
    \( \{ X - Y \} \)
    \( \{ X - Y \} \)
    \( \{ X - Y \} \)
    \( \{ X - Y \} \)
    \( \{ X - Y \} \)
    \( \{ X - Y \} \)

13. \( \{ X - Y \} \)
    \( \{ X - Y \} \)
    \( \{ X - Y \} \)
    \( \{ X - Y \} \)
    \( \{ X - Y \} \)
    \( \{ X - Y \} \)
    \( \{ X - Y \} \)
    \( \{ X - Y \} \)

14. \( \{ X - Y \} \)
    \( \{ X - Y \} \)
    \( \{ X - Y \} \)
    \( \{ X - Y \} \)
    \( \{ X - Y \} \)
    \( \{ X - Y \} \)
    \( \{ X - Y \} \)
    \( \{ X - Y \} \)

15. \( \{ X - Y \} \)
    \( \{ X - Y \} \)
    \( \{ X - Y \} \)
    \( \{ X - Y \} \)
    \( \{ X - Y \} \)
    \( \{ X - Y \} \)
    \( \{ X - Y \} \)
    \( \{ X - Y \} \)

16. \( \{ X - Y \} \)
    \( \{ X - Y \} \)
    \( \{ X - Y \} \)
    \( \{ X - Y \} \)
    \( \{ X - Y \} \)
    \( \{ X - Y \} \)
    \( \{ X - Y \} \)
    \( \{ X - Y \} \)

II. RESTRICTIONS ASSUMED IN B. (Cf. LASNIK & KUPIN 1977) BUT NOT IN A. (Cf. CHOMSKY 1955):

1. EVERY TERM A CONSTITUENT
2. 'ADJACENCY' CANNOT BE FREELY STIPULATED
3. LIMITED NO. OF TERMS (3)
4. ONLY ONE ELEMENTARY OPERATION PER T
5. ONLY 3 POSSIBLE ELEMENTARIES = SUBSTITUTION, (CHOMSKY) ADJUNCTION, PERMUTATION
6. I ASsume TO BE HEAVILY CONSTRAINED
7. NO 'TRAFFIC RULES' = ORDERING STATEMENTS, OPTIONAL V. OBLIGATORY RULES, RELATIVE OBLIGATORINESS.

II. WHY LOOK FOR RESTRICTIONS, AND WHAT RESTRICTIONS 'MAKE SENSE'?

1. THERE IS A (RELATIVELY LESS IMPORTANT) PRIMA FACIE ARGUMENT THAT U.C.
2. IS HEAVILY RESTRICTED = HEAVILY RESTRICTED THEORIES HAVE ACHIEVED A HIGH DEGREE
3. OF DESCRIPTIVE ACCURACY - THIS WOULD BE A MONUMENTAL COINCIDENCE ON ALTERNATIVE
4. ASSUMPTIONS
5. THE CENTRAL ARGUMENT IS THAT IT MUST BE POSSIBLE FOR THE LEARNER TO SELECT
6. THE CORRECT GRAMMAR BASED ON FUNDAMENTALLY LIMITED DATA. U.C. MUST BE SO
7. STRUCTURED THEN, THAT THE DATA DOES NOT UNDERDETERMINE THE CHOICE OF G.
20. We should expect to find restrictions that limit the class of G's that contribute to the smooth functioning of the evaluation metric and that obviate the need for generally unavailable data.

21. It above, eliminate the need for negative data, i.e., evidence that a string is ill-formed. The other properties in 14, at least severely restrict the class of 1's (hence, of G's). Further, they help assure that the G's used with primary data will be well scattered.

22. Certain constraints on the operation of rules can play a role also. Locality constraints guarantee that long 1's are not needed in the data.

23. Some constraints that don't obviously make sense:
   - No it may have precisely eleven terms! This wouldn't make a real dent in the class, and doesn't seem to follow from anything (unlike 13 above, which follows rather naturally from 14).
   - It's that move material to the left in English cannot apply in both NPs and G's to the extent that this is true, we would surely want it to follow from something.
   - Restriction to some inner circle in the chemistry hierarchy.
   - Note that the right theory couldn't be a circle in this picture at all: for example, no human language could be finite, but every circle allows G's for such lists.

24. Note that the goal of reducing the number of G's is completely orthogonal to this notion of 'power': a theory like 2 might have a very small number of G's (and they might be nicely scattered). Lasnik & Kipnis is such a theory.

25. Note also that a theory allowing, say, the entire class of C.F. G's would be virtually useless as a model of N.G. It would allow many G's with the wrong properties; would allow too many G's; would not obviously provide the needed scattering.

26. To the extent that restrictions are based on learnability considerations, we might expect them to be possibly over ridden under specific circumstances in which there is no learnability problem. The context term in 9, turns out to be one. Example: the theory of markedness provides other cases. Thus, there can be both optional and obligatory rules as long as obligatory is the unmarked case.