

LEARNABILITY IN LINGUISTIC THEORY

2/23/81

H. LASNIK

THE PROBLEM: WHAT IS THE NATURE OF L.A.D. SUCH THAT THE FOLLOWING PROCEDURE CAN TAKE PLACE?

1) DATA (D) → L.A.D. → GRAMMAR (G)

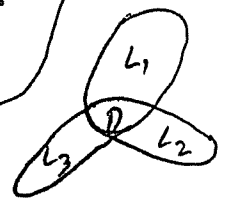
IF ANY LOGICALLY POSSIBLE GRAMMAR WERE AVAILABLE TO L.A.D., LANGUAGE ACQUISITION COULD NOT TAKE PLACE IN THE WAY THAT IT DOES. THE DATA WOULD DRASTICALLY UNDERDETERMINE THE GRAMMAR, AND A VARIETY OF CHILDREN COULD ARRIVE AT MUTUALLY INCONSISTENT G'S WHICH ALL AGREE ON THE PRESENTED DATA. THERE IS NO REASON TO BELIEVE THAT THIS HAPPENS.

2) D → LAO1 → G1      D → LAO2 → G2      D → LAO3 → G3 ...

WHERE G1 GENERATES L1      G2 ⇒ L2      G3 ⇒ L3 ...

AND L1 = D + E1      L2 = D + E2      L3 = D + E3 ...

AND E1, E2, E3, ... ARE ALL DIFFERENT.



MANY RESTRICTIONS ON WHAT COULD BE A (HUMAN) GRAMMAR HAVE BEEN HYPOTHESIZED. SEVERAL OF THE FOLLOWING HAVE MET WITH SOME EMPIRICAL SUCCESS:

3) - "X" RESTRICTIONS ON THE BASE COMPONENT

- LIMITATIONS ON THE NO. OF TERMS IN THE S.D. OF A T. (COMPARE SYNTACTIC STRUCTURES WITH LASNIK & KUPIN (1977))
- LIMITATIONS ON TYPE + NO. OF OPERATIONS A T CAN PERFORM
- ELIMINATION OF MANY 'TRAFFIC RULES' = ORDERING STATEMENTS, STIPULATIONS OF OBLIGATORINESS, AND THE LIKE

THESE ALL CLAIM THAT THE CLASS OF BIOLOGICALLY AVAILABLE G'S IS SMALLER THAN THE CLASS OF "MATHEMATICALLY AVAILABLE" G'S.

CERTAIN OTHER HYPOTHESES (ALONG WITH SOME OF THE ABOVE) SERVE TO EXPLAIN HOW ACQUISITION COULD TAKE PLACE GIVEN THE GENERAL LACK OF "NEGATIVE DATA" - i.e. INFORMATION THAT A PARTICULAR SENTENCE IS ILL-FORMED.

CERTAIN CASES OF ORDERING; OBLIGATORINESS; CONTEXT TERMS IN RULES ALL WOULD SEEM TO REQUIRE NEGATIVE EVIDENCE. INTERESTINGLY, THEY SEEM DISPENSABLE IN DESCRIBING SYNTACTIC FORM.

THERE ARE FURTHER ISSUES CONCERNING THE TYPE OF DATA AVAILABLE. IF SENTENCES WITH SEVERAL LAYERS OF EMBEDDING ARE NOT AVAILABLE IN THE DATA BASE, IT IS NOT SURPRISING THAT SYNTACTIC OPERATIONS <sup>(GENERALLY)</sup> SEEM TO OBEY 'LOCALITY' CONSTRAINTS.

IN THE PAST 10-15 YEARS, IT HAS BECOME MORE AND MORE APPARENT THAT PARTICULAR GRAMMARS HAVE LESS STRUCTURE THAN WE ORIGINALLY THOUGHT, AND THAT UNIVERSAL GRAMMAR (U.G.) HAS MORE STRUCTURE.

ONE WAY IN WHICH U.G. HAS A RICH STRUCTURE IS THE MODULARITY OF THE SYSTEM. THERE ARE SEVERAL COMPONENTS, EACH WITH ITS OWN PERMISSIBLE RULE TYPES AND PRINCIPLES. THIS PROVIDES FAR FEWER RULES (HENCE GRAMMARS) THAN AN UNDIFFERENTIATED THEORY SUCH AS POSTAL'S 'BEST THEORY'.

EVEN IF U.G. HAS 'FEW' GRAMMARS TO CHECK, THE SITUATION PICTURED IN (2) COULD STILL ARISE. IT WOULD ONLY FAIL TO ARISE IF U.G. WERE SO RESTRICTIVE THAT ONLY ONE G WERE ~~CONSISTENT~~ <sup>CONSISTENT</sup> WITH ~~ANY~~ ANY GIVEN D. THIS SEEMS UNLIKELY, HENCE THE THEORY OF MARKEDNESS (THE EVALUATION METRIC) WILL PLAY A ROLE.

NOT ONLY WILL U.G. MAKE AVAILABLE A CLASS OF GRAMMARS, THEN, BUT THESE G'S WILL BE (PARTIALLY) ORDERED IN TERMS OF PREFERENCE. THE G'S COMPATIBLE WITH D WILL BE RANKED, THEN. THE MOST HIGHLY VALUED ONE WILL BE TRIED FIRST, AND DISCARDED ONLY IF IT FAILS IN THE FACE OF NEW DATA.

SOME CONTRIBUTIONS TO THE THEORY OF MARKEDNESS :

- 4)
  - a. SYMBOL COUNTING - THE FEWER THE BETTER (CHOMSKY & HALLÉ)
  - b. S V.  $\bar{3}$  FOR SUBJACENCY (RIZZI AS INTERPRETED BY ANDERSON)
  - c. "UNIQUENESS" - WEXLER'S PROPOSAL THAT THE UNMARRIED CASE IS FOR ONLY ONE SURFACE STRUCTURE TO BE DERIVABLE FROM A GIVEN DEEP STRUCTURE.