CLARIFICATION

THE FIRST THING I'LL BE DISCUSSING IN THE SEMINAR IS CHOMSKY'S RECENT ARTICLE, WHICH WILL APPEAR IN THE MIT PRESS BOOK THAT ROGER MARTIN, DAVID MICHAELS, AND MYSELF ARE EDITING IN HONOR OF HOWARD LASNIK. SINCE I HAVE THE ELECTRONIC MANUSCRIPT, I'M TAKING THE LIBERTY OF COMMENTING ON IT, GLOSSING IT WITH MY OWN THOUGHTS (IN CAPITALS) AS WE GO ALONG. I KEPT THE FOOTNOTES AS THEY APPEAR IN THE TEXT, MAINLY BECAUSE I THINK IT'S USEFUL TO SEE THE DIGRESSIONS THAT ARISE IN THEM; I BREAK THE TEXT EVERY TIME A NOTE EMERGES. IT SHOULD BE CLEAR THAT, ALTHOUGH WE HAVE CHOMSKY'S PERMISSION TO DISCUSS HIS WORK IN THIS FORMAT, THIS IS NOT INTENDED FOR DISTRIBUTION AS SUCH.

<Minimalist Inquiries: the Framework>

The remarks that follow are "inquiries," a term intended to stress their tentative character. They are "minimalist" in the sense of the "Minimalist Program," itself exploratory as the term indicates, and in its short career already developing in partially conflicting and attractive directions.

THIS IS CRUCIAL TO KEEP IN MIND, AND IT IS WORTH EMPHASIZING THAT ARGUMENTS OF AUTHORITY ('CHOMSKY SAYS THAT...') ARE NOW MORE WORTHLESS THAN EVER.

What appears here is the first part of an unfinished manuscript.

I KNOW OF NO SECOND PART, UNFORTUNATELY, NOR WHAT CHOMSKY MEANS BY AN 'UNFINISHED MANUSCRIPT' (I.E. IN WHAT STAGE THAT WORK IS).

Here, I will keep to general considerations, rethinking the issues and concerns that motivate the program and attempting to give a clearer account and further development of them from one point of view, taking as a starting point the final sections of Chomsky (1995a) (henceforth MP>).

{As discussed in the Introduction of MP>, the chapters are largely based on lecture-seminars at MIT, the last of them in Fall 1994. What follows draws from discussions of Fall 1995, 1997. I will make no attempt to review the impressive range of recent work that bears directly on questions that arise, a failure that leaves no slight lacunae, as does the failure to consider alternatives that have been developed, among them, Abraham et al. (1996), Brody (1995), Collins (1997), Epstein et al. (1998), Frampton and Guttman (1998), Sportiche (1995), Zwart (1996). For comments on an earlier draft, many incorporated here, I am particularly indebted to Zeljko Boskovic, Chris Collins, Sam Epstein, Howard Lasnik, and Juan Uriagereka.}

That collection reflects an evolution over several years, with fairly radical changes along the way. Much like earlier stages, this one reflects a collective effort, incorporating ideas and proposals of many students and colleagues with no serious attempt at attribution, in fact no way to achieve it. I should, however, like to take the occasion to express my very special indebtedness to Howard Lasnik for many years of close collaboration, which has been extremely rewarding for me and is most inadequately recorded in print, though well-known to participants

in these enterprises.

<1. Background>

Let us begin by reviewing briefly a series of assumptions, discussed and qualified elsewhere. First, there is a faculty of language FL, a component of the human mind/brain dedicated to language. Given this endowment, a human infant, but not her pet kitten, will reflexively categorize parts of the confusion around her as "linguistic" and develop rich and highly articulated capacities to enter into these peculiar modes of human thought and action. In contrast, the infant and kitten will, it seems, develop along a rather similar path in acquiring capacities to deal with many other aspects of the world.

{See, e.g., Hermer and Spelke (1996). More generally, I assume that mental capacities are "modular" in the sense of Chomsky (1975), with "learning theories" LT(O,D) that may vary for organism O and cognitive domain D. The resulting modules might then have input/output properties of the kind analyzed in Fodor (1983), while belonging to a "central" system more structured than Fodor assumes.}

THAT LAST POINT IS AN INTERESTING DIVERGENCE WITH FODOR, WHICH HAS CONSEQUENCES TOUCHING ON THE COMPUTATIONAL THEORY OF MIND. INDEED, FODOR EXPLICITLY ARGUES THAT THE CENTRAL SYSTEM COULDN'T BE COMPUTATIONAL (AS RECENTLY AS IN HIS CONTRIBUTION TO THE "CHOMSKY CELEBRATION" A COUPLE OF MONTHS AGO), WHEREAS THE MODULES THEMSELVES ARE. THAT THE MODULES SHOULD BE COMPUTATIONAL, HOWEVER, AT LEAST IN A DEEP SENSE, IS FAR FROM OBVIOUS, AS JACKENDOFF HAS POINTED OUT. OF COURSE, THEY ARE SYSTEMATIC, BUT SO ARE CHEMICAL REACTIONS OR THE WORKINGS OF THE IMMUNE SYSTEM, AND WE DON'T (HAVE TO) CALL THEM COMPUTATIONAL. APART FROM SYSTEMATICITY, ONE WOULD ALSO NEED SOMETHING LIKE REPRESENTATIONALITY, WHICH IS NOT AN ALTOGETHER CLEAR NOTION, AT LEAST INTERNAL TO THE LINGUISTIC MODULE. (ALTERNATIVELY, OTHER COMPUTATIONAL NOTIONS, SUCH AS PHASES OF COMPUTATIONS, OR COMPUTATIONAL COMPLEXITY, MIGHT DETERMINE THE SERIOUS COMPUTATIONAL NATURE OF A SYSTEM.) EXTERNAL TO THE MODULE, SURELY THERE IS SOME INTENTIONAL CONNECTION BETWEEN LINGUISTIC ENTITIES AND WHAT THEY'RE USED TO REFER TO, AND ALSO (POSSIBLY) BETWEEN THOSE ENTITIES AND SOME 'MENTAL' PHENOMENA. NEEDLESS TO SAY, WHERE THE BOUNDARIES ARE HERE IS FAR FROM CLEAR, AND PART OF THE TASK OF MINIMALISM IS TO FIND OUT, PARTICULARLY SINCE WE TAKE BOUNDARY PROPERTIES (BARE OUTPUT CONDITIONS) TO GEAR SYNTAX. AT ANY RATE, IT IS POSSIBLE THAT, AFTER ALL, THE REPRESENTATIONAL CONNECTION LAY OUTSIDE OF THE GRAMMATICAL SYSTEM PROPER, PERHAPS OF FL ALTOGETHER. IF SO, IT COULD BE THAT WHAT'S SERIOUSLY COMPUTATIONAL ABOUT LANGUAGE IS NOT, CONTRA FODOR, IN THE MODULE ITSELF, BUT RATHER IN THE 'MORE STRUCTURED' CENTRAL SYSTEM. MUCH MORE ON THIS BELOW.

FL can be regarded as a "language organ," in the informal sense in which the visual system, or immune system, or circulatory system are commonly described as organs of the body: not

objects that can be be [SIC] removed leaving the rest intact, but subsystems of a more complex structure that we hope to understand by investigating parts that have distinctive characteristics, and their interactions.

THIS IS NOT AN INNOCENT REMARK. AN ORGAN IS ACTUALLY RATHER DIFFERENT FROM A MODULE. AS THE COMMENTS IMPLY, AN ORGAN IS A SUB-SYSTEM OF AN ORGANISM, WHEREAS A MODULE IS PROBABLY BEST CONCEIVED OF AS A PART OF AN ARTIFACT. YOU CAN REMOVE YOUR HARD DISK AND REPLACE IT, BUT IT ISN'T ENTIRELY CLEAR WHAT IT WOULD MEAN TO REPLACE YOUR CIRCULATORY SYSTEM. ALSO, THE HARD DISK OF YOUR COMPUTER WAS ASSEMBLED SEPARATELY FROM IT, WHEREAS YOUR CIRCULATORY SYSTEM GREW WITH YOU; CONVERSELY, YOUR COMPUTER MAY DIE, AND THE SCREEN OR KEYBOARD STILL BE USABLE, BUT WHEN YOU DIE YOUR CIRCULATORY SYSTEM GOES WITH YOU. GROWTH AND DECAY ARE NOTIONS THAT MAKE SENSE WITH ORGANS, BUT NOT SO CLEARLY WITH MODULES. SICKNESS, TOO, IS A PROPERTY OF ORGANS, AND PLASTICITY IN EXTREME CONDITIONS (WITHIN SOME POORLY UNDERSTOOD LIMITS, YOU MAY AID A MALFUNCTION IN SOME PARTICULAR ORGAN BY USING COMPONENTS OF OTHER ORGANS).

Despite explicit denials and apparent controversy, this much seems to be generally assumed, at least tacitly.

{For some illustrations and discussion, see Jenkins (1997, forthcoming), Marcus (1998).}

Like other organs, FL has an "initial state" S<0> that is an expression of the genes. To good first approximation, it is uniform for the species, apparently also biologically isolated in essential respects

THIS ARE ALL INDICATIONS THAT THE NOTION 'ORGAN' IS NOT TAKEN METAPHORICALLY.

and a very recent evolutionary development.

{Which implies virtually nothing about the novelty of its component elements.}

AGAIN, AS EXPECTED FROM ORGANS. FOR INSTANCE, THE BONES OF YOUR MIDDLE EAR ARE THE RESULT OF GROWTH AND STRUCTURAL CHANGES IN THE REPTILIAN CRANIUM, WHERE THEY WERE PART OF THE JAW, AND AS SUCH A MUCH OLDER COMPONENT OF THE RELATIVELY NEW MAMMALIAN HEARING SYSTEM. FL COULD HAVE HAD A LOT OF OLD SUBCOMPONENTS THAT NEVER GOT TO BE ASSEMBLED INTO WHAT WE NOW UNDERSTAND AS LANGUAGE UNTIL SOME EVOLUTIONARY EVENT LINKED THEM ALL, 'APPROPRIATELY' FROM OUR PRESENT PERSPECTIVE. IN OTHER WORDS, THERE'S NO REASON TO BELIEVE ANY GRADUALISM IN THE PROCESS, OR FOR THAT MATTER THAT EVERYTHING FELL INTO PLACE RECENTLY; WE JUST DON'T KNOW AND THE LOGIC OF EVOLUTION DOESN'T HELP DECIDE.

FL undergoes state changes under triggering and shaping influences of the environment.

IT IS VERY IMPORTANT THAT THIS IS PHRASED THIS WAY, SINCE IT POINTS TOWARDS A RELATIVELY DEEP THEORY OF WHAT A PARAMETER IS. THE OLD IDEA THAT GENES ARE RECIPES FOR INDIVIDUALS IS LONG GONE. NO GENE IS RESPONSIBLE FOR YOUR HAND OR YOUR FACE. THESE COMPLEX STRUCTURES ARE THE RESULT OF AN INTRICATE INTERPLAY BETWEEN MASTER CONTROL GENES AND ELEMENTS IN THE EARLY ENVIRONMENT (IN THIS CASE, IN UTERO). WITHOUT BOTH GENES AND THE ENVIRONMENT YOU GET NOTHING REMOTELY CLOSE TO AN INDIVIDUAL AS YOU NOW SEE IT. THERE'S NO REASON WHY THESE INTERACTIONS SHOULD TAKE PLACE JUST IN UTERO. INDEED, MANY CREATURES GET TO BE STRUCTURALLY DETERMINED ONLY AFTER EARLY ENCOUNTERS WITH SOMETHING IN THE ENVIRONMENT (FOOD, A POTENTIAL MATE, LIGHT, NOISE, TEMPERATURE, ETC.). FOR INSTANCE, SEX IS OFTEN DETERMINED IN TERMS OF OUTSIDE CONDITIONS. IF THIS IS THE CASE FOR STRUCTURE, THERE'S EVERY REASON TO BELIEVE IT IS ALSO THE CASE FOR BEHAVIOR. THUS, FOR INSTANCE, BIRD SONGS WHICH GET FIXED ONE WAY OR THE OTHER IN TERMS OF EARLY EXPOSURE TO OTHER BIRD SONGS (SCIENTISTS ARE NOW BEGINNING TO UNDERSTAND THE CHEMICAL BASE OF SUCH SYSTEMS). FROM THIS PERSPECTIVE, THERE'S NOTHING PARTICULARLY DEEP ABOUT HAVING 'STATE CHANGES' (SWITCHES FROM THE INITIAL CONDITION OF FL TO SOME ADULT CONDITION) ON THE BASIS OF 'TRIGGERING AND SHAPING INFLUENCES OF THE [LINGUISTIC] ENVIRONMENT'. OF COURSE, YOU CAN ASK WHY THIS IS, BUT THE QUESTION EXTENDS TO OTHER CREATURES AND OTHER KINDS OF BEHAVIORS OR STRUCTURES. IT SEEMS AS IF NATURE IS ECONOMIC ENOUGH NOT TO SPECIFY IN THE GENES WHAT THE ENVIRONMENT DOES GIVE YOU FOR FREE IN A SYSTEMATIC FASHION. (ANOTHER WAY OF SAYING THAT IS THAT THERE IS NO SELECTIVE PRESSURE IN SUBSUMING A CHARACTERISTIC X IF THE ENVIRONMENT DOESN'T PARTICULARLY DEMAND IT OR FAVOR IT, IN THAT THE ENVIRONMENT ALREADY GRANTS YOU THAT CHARACTERISTIC X.) NOTE THAT IF THIS WERE THE CASE, THE INITIAL STATE OF FL WOULDN'T EVEN REMOTELY LOOK LIKE A LANGUAGE, LESS SO A SO-CALLED FORMAL LANGUAGE, JUST LIKE THE SYSTEM OF CONTROLLING GENES THAT RESULT IN A HAND DOESN'T EVEN REMOTELY RESEMBLE THE STRUCTURE OF A HAND, TRIVIALLY.

If Jones's FL is in state L, we say that Jones has (speaks, knows,...)

THIS OFTEN GETS PHILOSOPHERS GOING, PARTICULARLY THE USE OF THE TERM 'KNOW'. OF COURSE, THE INTENTION IS SOMETHING LIKE 'BEING ENDOWED WITH STRUCTURAL PROPERTY FL', MUCH AS ONE WOULD SAY THAT JONES IS ENDOWED WITH WHATEVER STRUCTURAL PROPERTY IS INVOLVED IN HIS IMMUNE SYSTEM, SAY. HARASSMENT OVER THE TERM 'KNOW' SEEMS A BIT OUT OF PLACE, PARTICULARLY BECAUSE NOBODY SEEMS TO HAVE A VERY CLEAR PICTURE OF WHAT (OTHER KINDS OF) KNOWING REALLY IS. AT ANY RATE, LINGUISTS WOULDN'T HAVE TO YIELD MUCH IF THEY GRANT THE PHILOSOPHICAL POINT, AND INSTEAD USE THE TERM 'COGNIZE' OR ANY OTHER FAVORITE; THE INTENTION BEHIND THE TERM IS CLEAR. AND OF COURSE THE ISSUE OF THE REPRESENTATION OF LINGUISTIC ENTITIES (HOW OR EVEN

WHETHER THE STRUCTURAL PROPERTIES AND INTERACTIONS THAT LINGUISTS STUDY RELATE TO GOD-KNOWS-WHAT 'MIND' ENTITIES), FOR SOME PHILOSOPHERS A PREREQUISITE FOR SPEAKING ABOUT KNOWLEDGE HERE, IS UP FOR GRABS, AS MENTIONED ABOVE AND AS WILL BE CLEAR BELOW.

the (I-)language L.

{I assume here familiar idealizations, abstracting from real-world interactions that yield complex and widely varying forms of multiple systems. The term "idealization" sometimes misleads; the process is a crucial part of the effort to determine reality.}

THE IDEA HERE IS TO STAY AWAY FROM THE PHILOSOPHICAL NOTION OF LANGUAGE, OR FROM FORMAL LANGUAGES, OR ANY SUCH SYSTEM THAT INVOLVES MORE OR LESS DIRECT RELATIONS BETWEEN THE LINGUISTIC SYSTEM AND THE WORLD. THOSE EXIST, OBVIOUSLY, AND ARE UNDERSTOOD TO BE PART OF THE PERFORMANCE SYSTEMS (THUS NOT JUST 'SEMANTIC' RELATIONS, BUT ALSO 'PHONETIC' RELATIONS WITH THE WORLD, IN TERMS OF THE PHYSICS OF SPEECH AND SO ON). I-LANGUAGE IS MEANT IN THE SENSE OF THE FAMILIAR SYSTEM, INTERNAL TO SPEAKERS MINDS, INTENTIONALLY CHARACTERIZED, APPROPRIATELY IDEALIZED, THAT LINGUISTS OF OUR TRADITION STUDY.

Two immediate tasks of a theory of language are to characterize the languages (states) attained and the shared initial state: the tasks of "descriptive adequacy" and "explanatory adequacy," respectively.

DESCRIPTIVE AND EXPLANATORY ADEQUACY ARE PRESENTED HERE IN SLIGHTLY LESS METHODOLOGICAL TERMS THAN A FEW YEARS AGO, WHEN DESCRIPTIVE ADEQUACY REFERRED TO A CLASS OF THEORIES THAT CAPTURED THE OBSERVED FACTS, AND EXPLANATORY ADEQUACY TO A MODEL FOR DECIDING AMONG THOSE COMPETING THEORIES--IN EFFECT, FOR A THEORY OF ACQUISITION. THE PRINCIPLES AND PARAMETERS MODEL FOCUSED THINGS IN TERMS OF A GIVEN ONTOLOGICAL PICTURE, PRESUMED CORRECT (GIVEN THE EXTRAORDINARY DESCRIPTIVE SUCCESS OF THAT THEORY NOT JUST IN TERMS OF LANGUAGE DESCRIPTION, BUT ALSO BY ANY RATIONAL METRIC: PROGRESS IN PSYCHOLINGUISTIC EXPLANATIONS, BOTH IN ACQUISITION, MATURATION, AND PARSING, SOME CONCEPTIONS OF NEUROLINGUISTIC SUPPORT AND PREDICTIONS FOR LANGUAGE BREAKDOWN, THEORIES OF LANGUAGE CHANGE, AND SO ON). WITHIN THAT PICTURE, EXPLANATORY ADEQUACY REDUCES (!) TO UNDERSTANDING THE DETAILS OF THE INITIAL STATE OF THE FACULTY, WHEREAS DESCRIPTIVE ADEQUACY IS INTENDED IN THE SENSE OF SHOWING US HOW TO MAP THAT SYSTEM TO WHATEVER IT IS THAT SPEAKERS KNOW, COGNIZE, OR WHATEVER IT IS THEY DO. THESE ARE TAKEN TO BE 'IMMEDIATE' TASKS; MORE REMOTE ONES, BUT EQUALLY IMPORTANT IN THE END, ARE THE RAMIFICATIONS (EXACTLY HOW THIS SYSTEM IS PUT TO USE, EXACTLY HOW IT IS FOUND IN THE BRAIN, AND SO ON).

We understand Universal Grammar (UG) to be the theory of the initial state, and particular

grammars to be theories of attained states.

IT SHOULD BE EMPHASIZED THAT UG IS A THEORY OF AN ORGAN, NOT ITSELF AN ORGAN (FL). SIMILARLY, THE I-GRAMMAR OF ENGLISH IS IN EFFECT A THEORY OF HOW FL GOT FROM ITS INITIAL STATE TO WHATEVER STATE IT HAS IN OUR MINDS RIGHT NOW, APPROPRIATELY IDEALIZED.

The language L includes a cognitive system that stores information: roughly, information about sound, meaning, and structural organization.

COGNITIVE SYSTEM STANDS HERE FOR THE OLD 'COMPETENCE'. IT IS A MORE APPROPRIATE TERM, SINCE PHILOSOPHERS, AGAIN, COMPLAIN ABOUT 'COMPETENCE' BEING USED FOR OTHER NOTIONS (A SHOE-MAKER BEING COMPETENT IN SHOE-MAKING AFTER HAVING LEARNED THE CRAFT). AGAIN, IT SHOULD BE CLEAR WHAT IS, IN THIS CASE INNOCENTLY, MEANT.

Performance systems access this information and put it to use.

THIS IS NON-TRIVIAL, OF COURSE. MANY COMPLAIN HERE THAT THIS SHOULDN'T BE THE PICTURE OF PERFORMANCE SYSTEMS AT ALL, THAT THOSE SYSTEMS IS WHAT IS REAL NOW AND HERE, AND THE COGNITIVE SYSTEMS ARE MERE IDEALIZATIONS (THE TERM USED PEJORATIVELY TO MEAN FIGMENTS OF THE THEORIST'S IMAGINATION). NEEDLESS TO SAY, THAT COULD BE TRUE, BUT THE IMPORTANT THING TO KEEP IN MIND HERE IS THAT THIS IS A SCIENTIFIC HYPOTHESIS, SO THERE IS NO SENSE IN CHASTISING IT A PRIORI, WITH SOME DOGMATIC POSTURING. THE HYPOTHESIS BEING ENTERTAINED (RIGHT OR WRONG AS ANY OTHER EMPIRICAL CLAIM) MAKES PERFECT SENSE. EVOLUTION ENDOWED HUMANS WITH A CERTAIN STRUCTURAL PROPERTY FL, JUST AS IT ENDOWED OTHER SPECIES (INCLUDING HUMANS) WITH PROPERTIES SUCH AS THOSE INVOLVED IN THE PLAN OF THEIR ORGANS, FOR INSTANCE THOSE INVOLVED IN LOCOMOTIVE SYSTEMS. DOES IT MAKE SENSE TO SAY THAT PERFORMANCE SYSTEMS FOR LOCOMOTION ACCESS SOME KIND OF STRUCTURAL INFORMATION AND PUT IT TO USE? SURELY THAT MUST BE GOING ON WHEN OUR LOCOMOTIVE ORGANS GET US TO WALK, JUMP, CRAWL, SWIM, OR WHATEVER, AND IN OUR CASE NOT TO FLY, GLIDE, OR EVEN GALLOP, AND SO FORTH. IT IS EASY TO SEPARATE THE LOCOMOTIVE STRUCTURAL SYSTEM FROM ITS BEING PUT TO USE; INDEED, PEOPLE WORKING ON LOCOMOTION ACROSS SPECIES HAVE FOUND ONLY A FEW ANATOMICALLY VIABLE WAYS THAT ESSENTIALLY ALL STUDIED SPECIES (FROM INSECTS TO REPTILES) HAVE ACCESS TO, FOR WHATEVER REASON (PHYSICAL CONSTRAINTS, MATHEMATICAL WAYS OF RHYTHMICALLY RELATING SEVERAL LEGS, ETC.). THERE IS A REAL SENSE IN WHICH ONE COULD SPEAK OF THE COGNITIVE/ANATOMICAL SYSTEM FOR LOCOMOTION IN ANIMALS, AND THEN OF HOW THAT SYSTEM GETS TO BE DEPLOYED IN GIVEN SPECIES, AND INDEED EVEN GIVEN INDIVIDUALS DEPENDING ON VARIOUS FACTORS (ONE CAN IMAGINE SLIGHTLY DIFFERENT LOCOMOTIVE SYSTEMS BEING FIXED IN GIVEN INDIVIDUALS DEPENDING ON VARIOUS ENVIRONMENTAL CONDITIONS). ONCE THAT'S ACCEPTED, IT'S NOT PARTICULARLY DIFFICULT TO ACCEPT THE CLAIM

THAT PERFORMANCE SYSTEMS IN INDIVIDUALS (FOR INSTANCE, THOSE CONNECTING THE GENERAL LOCOMOTIVE SYSTEM AS DEPLOYED IN A GIVEN INDIVIDUAL TO ITS RESPIRATORY SYSTEM, PERHAPS TO SOME BALANCING SYSTEM, OR EVEN TO SYSTEMS RESPONSIBLE FOR KEEPING RHYTHMIC PATTERNS IN SYNCH, AND SURELY MEMORY SYSTEMS RESPONSIBLE FOR STORING CERTAIN MOVES IN GIVEN TERRAINS, AND SO ON) ACCESS THE BASIC, 'STRUCTURALLY STORED' INFORMATION AND PUT IT TO USE. THIS COULD BE WRONG, OF COURSE, BUT IT ISN'T IMPOSSIBLE OR EVEN PARTICULARLY INSANE.

Empirical questions arise at once: in particular, to what extent are the performance systems part of FL, that is, language-dedicated, specifically adapted for language?

{As systems, that is; their components need not be. See note

THIS MAKES REFERENCE TO THE NOTE ABOUT PARTS OF SYSTEMS EVOLVING INDEPENDENTLY OF THE WHOLE, PERHAPS WITH DIFFERENT PURPOSES. THE KEY EMPIRICAL QUESTION IS WHETHER PERFORMANCE SYSTEMS AS A WHOLE ARE DEDICATED TO LANGUAGE, OR SOMEHOW 'CO-OPTED' FOR LINGUISTIC USE IN THE PERFORMANCE INSTANCE. AGAIN, THE WHOLE POINT HERE IS THAT WE CANNOT DECIDE THIS A PRIORI, OBVIOUSLY. NOTICE THE SUBTEXT. FOR ADAPTATIONISTS, THE ANSWER IS GIVEN A PRIORI: HOW COULD YOU HAVE A PERFORMANCE SYSTEM THAT IS NOT PERFECTLY ADAPTED TO LANGUAGE USE? CHOMSKY (WITH GOULD AND MANY OTHER CRITICS OF THE NEO-DARWINIAN SYNTHESIS) THE QUESTION IS THE OPPOSITE: WHY SHOULD IT BE THAT THE ADAPTATION IS PERFECT? WHAT WOULD HAVE GONE WRONG WITH THE LOGIC OF EVOLUTION IF THE ADAPTATION HAD NOT BEEN PERFECT? NOTHING MUCH. IT'S NOT AS IF WHATEVER SPECIES GOT THE BENEFIT OF BEING ABLE TO 'MORE OR LESS' USE FL WOULD BECOME EXTINCT JUST BECAUSE IT DIDN'T GET TO USE IT 'COMPLETELY'. YOU GET WHAT YOU GET, AND OTHERS OF YOUR SPECIES GET THE SAME AS YOU DO, SO YOU ADAPT WITHIN THOSE PARAMETERS, AND SO LONG AS THE RESULTING STATE DOESN'T KILL YOU BEFORE TRANSMITTING YOUR GENES, THERE'S ABSOLUTELY NOTHING THAT SHOULD KEEP YOU FROM THRIVING IN YOUR LUCKY EVOLUTIONARY FATE.

<NOV>.} On the "sound side," the answer is unclear and disputed;

WHAT'S AT ISSUE, AT LEAST IN PART, HERE IS WHETHER, SAY, SIGNED LANGUAGES, WHICH OBVIOUSLY DO NOT INVOLVE SOUND SYSTEMS, NONETHELESS ARE EQUALLY ATTUNED WITH FL IN TERMS OF WHATEVER GESTURES THEY USE. IF PERLMUTTER IS TO BE BELIEVED, PHONOLOGY (WHATEVER THAT IS) TURNS OUT TO BE WAY MORE ABSTRACT THAN ANYONE IMAGINED, WITH HAND POSITIONS BEING THE EQUIVALENT OF CONSONANTS AND HAND MOVEMENTS OF VOWELS. IF SO, EITHER THE 'REAL' (SOUND) PHONETIC SYSTEM IS NOT CRUCIALLY LANGUAGE-DEPENDENT, OR ELSE IT IS JUST A SURFACE MANIFESTATION OF A DEEPER SYSTEM INCLUDING WHATEVER IS GOING ON IN SIGN LANGUAGE. AT THE SAME TIME, LOTS OF EXPERIMENTS EXIST INDICATING THAT THERE IS SOMETHING SPECIFICALLY LINGUISTIC ABOUT PHONETIC SYSTEMS (AS OPPOSED TO GENERAL SOUND), WHICH

SUGGESTS THE SECOND OF THOSE OPTIONS IS CORRECT.

on the "meaning side," the questions are much harder and more obscure for obvious reasons, and judgments can only be highly tentative.

SURPRISINGLY, THIS IS HARDLY EVER RECOGNIZED BY SEMANTICISTS OR PHILOSOPHERS, WHO SHOULD KNOW BETTER. IN ALL DUE RESPECT, 'WHAT FREGE SAID' OR 'WHAT RUSSELL BELIEVED' ARE NOT STATEMENTS WE WOULD TOLERATE ON THE PHONETIC SIDE (THAT IS, WHAT CONVINCES PEOPLE THERE IS HOW AN EXPERIMENT TURNS OUT OR HOW TIGHT A THEORY IS AND HOW WELL IT ACCOUNTS FOR THE FACTS, NOT WHETHER 'TRUBETZKOY SAID IT', OR SOME SUCH AUTHORITY ARGUMENT).

A standard working assumption is that performance systems are external to FL. That is a simplifying assumption, not definitely known to be false, though it may well be, perhaps in important ways.

{Processing systems vary with languages and language types, even for very young infants, enabling them to sort out distinct languages in the data to which they are exposed. See Bosch and Sebastián-Gallés (1997), Jusczyk (1997), Mehler and Dupoux (1994). Whether these important discoveries (which add new dimensions to "poverty of stimulus" arguments) are consistent with the simplifying assumption depends on how sensorimotor processing is "modulated" by the target language.

WHATEVER THAT MEANS, THOUGH. IT IS HARD TO SEE HOW THIS MODULATION COULD EVER ARISE, ALTHOUGH WHO KNOWS. NOTE ALSO BELOW:

Whorfian ideas on the meaning side have a similar flavor.

INDEED, THIS IS THE CASE, AND SIMILAR SKEPTICISM IS POSSIBLY IN ORDER. THAT IS, HOW DOES A CHILD ACQUIRE A PARAMETRIC DIFFERENCE IN THE SEMANTIC SIDE? WHAT KIND OF EVIDENCE TRIGGERS THAT PARAMETER? PERHAPS IT EXISTS, BUT I HAVEN'T SEEN A SINGLE CASE THAT SHOWS HOW A GIVEN COMMUNICATIVE SITUATION THAT THE CHILD CAN FACE WOULD LEAD TO THIS OR THAT SEMANTIC PARAMETRIC SETTING.

See Phillips (1996), for an intriguing approach that bears on some of these questions.}

Theissues have had little effect on empirical inquiry into questions of descriptive and explanatory adequacy, but come to the fore within the minimalist program.

THE REASON THEY DO IS TWOFOLD. BROADLY, MINIMALISM IS AN EXERCISE IN BOUNDARIES AND--IF YOU WISH, THE BIG PICTURE--SO YOU MUST WORRY ABOUT HOW TO ADDRESS THESE QUESTIONS, AND WHAT THEY TELL YOU ABOUT WHERE THE THEORETICAL FRONTIERS BETWEEN THIS AND THAT ARE. MORE NARROWLY, MINIMALISM WORKS WITH THE HYPOTHESIS THAT INTERFACE CONDITIONS (READ AS 'INTERACTIONS WITH PERFORMANCE SYSTEMS') DRIVE THE SYNTACTIC ENGINE. IF SO, IT IS OF PARAMOUNT IMPORTANCE TO DETERMINE WHETHER THE INTERFACE IS SPECIFICALLY

LINGUISTIC OR NOT. IF IT IS YOU EXPECT LESS WILD THINGS TO HAPPEN THAN IF IT IS NOT, ALTHOUGH AT THE SAME TIME YOU WILL HAVE LESS AID FROM THE OUTSIDE SYSTEMS. THAT IS, IF PF IS PURELY (OR MOSTLY) LINGUISTIC, GENERAL THEORIES OF SOUND, ARTICULATION, OR WHATEVER MAY HAVE VERY LITTLE BEARING ON ITS SHAPE, AND THUS ON HOW PF AFFECTS A SYNTACTIC COMPUTATION. SIMILARLY FOR LF, THOUGH MORE OBSCURELY THERE: IF LF IS PURELY (OR MOSTLY) LINGUISTIC, GENERAL THEORIES OF MEANING, INFORMATION, CONCEPTUALIZATION, AND SO FORTH, MAY ESSENTIALLY BE IRRELEVANT, OR SO REMOTE FROM EVERYTHING THAT WE MAY AS WELL PROCEED WITH OUR STANDARD STUDY OF LF, THAT IS SEEKING LINGUISTIC PATTERNS AND LEAVING THEIR SEMANTIC CONNECTION FOR LATER. I'M NOT SAYING ANY OF THIS SKEPTICISM IS FOUNDED, JUST THAT IT DEPENDS ON ANSWER THE QUESTION OF SPECIFICITY AT THE INTERFACE.

I will adopt usual conventions for present purposes, recognizing that they are not innocent. We therefore take L to be a cognitive system alone.

IT IS CLEAR WHAT POSITION CHOMSKY'S BETTING ON, FOR BETTER OR FOR WORSE.

I will assume further that L provides information to the performance systems in the form of "levels of representation," in the technical sense.

{I am assuming here the basic framework of Chomsky (1955-6), though of course there have been radical changes since. Levels are systems of representations; representations formed in the course of derivation typically do not form part of a level.

THIS IS A CRYPTIC COMMENT, BUT IT IS EASILY UNDERSTOOD IF YOU GO BACK TO THE EARLIER SYSTEM, OR EVEN ASPECTS. A LEVEL OF REPRESENTATION IS IN ESSENCE A UNIFIED OBJECT WITH SUB-COMPONENTS BUILT ON THE BASIS OF SPECIFIC INTERACTIONS AND SYMBOLS FROM A VOCABULARY OF PRIMES. CLASSICALLY, D-STRUCTURE, OR EVEN BEFORE THAT, THE P-MARKER, WERE THOUGHT OF AS LEVELS, IN THIS INSTANCE RESULTING FROM THE INTERACTION OF PHRASE STRUCTURE RULES (SAY) ON CERTAIN SYMBOLS, IN FAMILIAR WAYS. THE T-MARKER OF EARLIER MODELS, OR S-STRUCTURE AND LF OF LATER ONES, RESULTED FROM MODIFYING THE INITIAL LEVEL BY WAY OF TRANSFORMATIONS. LATER MODIFICATIONS TOOK S-STRUCTURE OR THE T-MARKER TO THE MORPHO-PHONOLOGICAL LEVEL, OR PF, AND SO ON. NATURALLY, YOU NEED FUNCTIONS TO CARRY YOU FROM ONE LEVEL TO THE NEXT, AT LEAST IN A DERIVATIONAL SYSTEM. THAT WAS THE ROLE OF TRANSFORMATIONS, CLASSICALLY. IT IS PERHAPS USEFUL TO SEPARATE THE LEVEL OF S-STRUCTURE, LF, OR WHATEVER, FROM THE COMPONENT OF LF, OR THE TRANSFORMATIONAL COMPONENT. A COMPONENT IS A COLLECTION OF OPERATIONS THAT YIELD PARTIAL REPRESENTATIONS, WHICH EVENTUALLY ARRAY THEMSELVES INTO A LEVEL. ALTHOUGH THEY DON'T HAVE TO, MIND YOU. CHOMSKY IS GOING WITH A SYSTEM WHERE THE PERFORMANCE SYSTEM, BY ARCHITECTURAL FIAT, ACCESSES A LEVEL, THAT IS ACCESSES FL AT A GIVEN, SINGLE POINT. HE IS THEREBY NOT PURSUING AN ALTERNATIVE IN

WHICH THE PERFORMANCE SYSTEM ACCESSES THE PARTIAL REPRESENTATIONS THAT ARE FORMED IN THE SYSTEM AS THE COMPONENTS UNFOLD. YOU HAVE TO REALIZE THIS IS AN EMPIRICAL DECISION, ALTHOUGH CHOMSKY HASN'T GIVEN US (AT LEAST HERE) ANY ARGUMENTS IN FAVOR OF THE SYSTEM CONVERGING INTO LEVELS, AS OPPOSED TO PERFORMANCE ACCESSING THE COMPONENTS DIRECTLY.

Note further that the term "representation" is a technical one, with no "representation" relation in the sense of representational theories of ideas, for example.}

PHILOSOPHERS WILL SURELY ARGUE THAT THIS IS NOT SUCH A SIMPLE, TERMINOLOGICAL MATTER. AT ISSUE IS WHETHER THE, IF YOU WISH, 'TECHNICAL REPRESENTATIONS' IN CHOMSKY'S SENSE CONNECT WITH SOMETHING, EITHER 'OUT THERE' IN THE WORLD, OR 'IN HERE' WITHIN THE MIND/BRAIN. THAT IS, FOR INSTANCE, THE LF-LEVEL FOR ANY OF THESE PHRASES, OR THE PF-LEVEL, OR IF YOU WISH EVEN THE SYMBOL 'NOUN PHRASE' OR 'P' OR WHATEVER. IS THAT SYMBOL A 'REPRESENTATION', IN THE SENSE THAT IT CORRESPONDS TO SOMETHING, SAY, IN THE MIND? IF NOT, WHAT IS IT? IF SO, HOW DOES IT CORRESPOND? CHOMSKY HAS PURPOSELY NOT WANTED TO ADDRESS THIS ISSUE, SINCE HE THINKS IT'S EITHER PREMATURE TO ASK THAT OR IT MAY EVEN BE AN ILL-POSED QUESTION. IF WE WERE ASKING IT ABOUT THE IMMUNE SYSTEM, SAY, WOULD IT MAKE SENSE? AND EVEN IF IT DID ('HOW ARE VIRUSES REPRESENTED IN OUR ORGANISM?'), WOULD WE STOP DOING IMMUNOLOGY BECAUSE WE DIDN'T QUITE KNOW WHETHER THERE IS AN INVENTORY OF SIGNALS IN OUR ORGANISM FOR VIRUSES 'KNOWN' TO OUR SPECIES, OR RATHER GIVEN INDIVIDUALS GET TO 'KNOW' VIRUSES AS THEY ARE EXPOSED TO THEM? THE ANSWER HERE IS CLEAR. BY PARITY OF REASON, THE QUESTION OF LINGUISTIC REPRESENTATIONS IN THE PHILOSOPHICAL SENSE IS TOO VAGUE TO ADDRESS VERY MEANINGFULLY. IF OURS ARE, AFTER ALL, NOT REPRESENTATIONS IN ANY 'SERIOUS' SENSE OF THE WORD, WELL, SO BE IT--THEY'RE NOT. THAT WON'T AFFECT ANYTHING IN LINGUISTICS (ALTHOUGH IT MIGHT HAVE SOMETHING TO SAY ABOUT THE THEORY OF MIND, UNFORTUNATELY NOT VERY PROMISING: IF THE THEORY OF MIND GETS RID OF THE SUBFIELD WHERE MOST PROGRESS HAS BEEN MADE BECAUSE IT DOESN'T CONFORM TO THE REPRESENTATIONAL CREDO, WELL, GOOD LUCK TO THE THEORY OF MIND IN ITS WORKS WITH LOBSTERS' GUT FEELINGS AND THE PERIPHERAL VISION OF FRUIT FLIES).

The performance systems access these "interface levels." Assume further that performance systems are of two kinds: sensorimotor systems and systems of thought (to give a name to something very poorly understood). Let us take them (tentatively) to be unitary and distinct, in the sense that all sensorimotor systems access one interface level, and all systems of thought access a distinct interface level.

IT SHOULD BE CLEAR THAT THESE ASSUMPTIONS ARE WILDLY TENTATIVE. THAT THERE IS A SINGLE POINT OF ACCESS TO THE SENSORIMOTOR SYSTEM HAS SOMETHING GOING FOR IT, ALTHOUGH HIGHLY CONTROVERSIAL: THE

MOTOR THEORY OF SPEECH PERCEPTION. A SIMILAR BELIEF ABOUT THE 'MEANING' SIDE IS, AS FAR AS I CAN SEE, A MERE RESIDUE OF ARISTOTELIAN THOUGHT, TURNED INTO DOGMA BY TRADITION. THERE IS NO IOTA OF EVIDENCE TELLING US THAT 'THOUGHT' (INDEED VERY POORLY UNDERSTOOD) ACCESSES FL AT A SINGLE POINT. TO THE EXTENT WE EVEN REMOTELY UNDERSTAND WHAT'S AT ISSUE AT THE POINT OF ACCESS, WE KNOW THIS HAS TO DO WITH (AT LEAST) TWO THINGS. ONE IS CONCEPTUAL INFORMATION, OF THE SORT SEPARATING NOUNS FROM VERBS, AND AMONG THE LATTER COUNT AND MASS AND SO ON, AND AMONG THE FORMER STATES AND EVENTS AND SO ON, AND OF COURSE WAYS OF LIMITING NAMING, WAYS OF LIMITING NOTIONALLY POSSIBLE ARGUMENTS (THEMES, AGENTS, GOALS, ETC.) AND SURELY OTHERS. ANOTHER, A PRIORI TOTALLY DIFFERENT SYSTEM INVOLVES INTENTIONAL INFORMATION, OF THE SORT INVOLVED IN REFERRING, QUANTIFYING, GETTING SCOPAL INTERACTIONS, CONTEXTUAL RESTRICTIONS ON PREDICATION, AND SO ON. THE REASON I'M SAYING THESE ARE DIFFERENT IS EMPIRICAL. NO KNOWN INTENTIONAL PHENOMENON AFFECTS CONCEPTUAL PRESENTATION (NO LANGUAGE, FOR INSTANCE, USES A SPECIAL VERB FOR A GIVEN SCOPAL INTERACTION, OR A PARTICULAR NOUN IF IT HAPPENS NOT TO HAVE REFERENCE IN THE REAL WORLD, AND SO ON); AND VICE-VERSA, NO KNOWN CONCEPTUAL PHENOMENON AFFECTS INTENTIONAL STANCES (NO LANGUAGE HAS LIMITATIONS ON WHAT CAN BE A PREDICATE GIVEN CONCEPTUAL DIFFERENCES, WHETHER SOMETHING IS A STATE, AN EVENT, AND SO FORTH, OR EVEN AN INDIVIDUAL FOR THAT MATTER). THESE JUST SEEM UNRELATED, AND WERE TAKEN TO BE UNRELATED IN ALL MODELS THAT SERIOUSLY BOTHERED TO EXPLORE THEM (E.G. D-STRUCTURE VS. LF). NOW WE'RE ASSUMING THAT CONCEPTUAL STRUCTURE AND INTENTIONAL STRUCTURE GO INTO THE SAME LEVEL. WE WILL RETURN TO THIS ASSUMPTION LATER ON, AND WHY IT IS MADE. AT THIS POINT I JUST WANT TO EMPHASIZE THE FIRST POINT WHERE IT IS ACKNOWLEDGED THAT THERE IS NOT PARTICULARLY GOOD REASON TO MAKE IT. IN OTHER WORDS: BEWARE, SINCE THIS CAN BE A PART OF THE PROGRAM THAT CAN CHANGE, AS WE UNDERSTAND WHETHER OR NOT WE SHOULD REALLY MAKE THIS ASSUMPTION.

On these assumptions we understand L to be a device that generates expressions EXP = <PHON, SEM>, where PHON provides the "instructions" for sensorimotor systems and SEM for systems of thought; information about sound and meaning, respectively, where "sound" and "meaning" are understood in internalist terms, "externalizable" for language use by the performance systems.

{On my own views on the issues, see Chomsky (1975), (1995b), (1996), among others.}

Theories of PF and LF seek to spell out the nature of PHON and SEM. I will assume some version of standard theories to be adequate for present purposes, using the conventional term "features" for the properties of language that enter into PF, LF, and the computational system that generates them.

THIS ISN'T INNOCENT EITHER, CONVENTIONAL THOUGH THE TERM 'FEATURE' IS.

IT HAS THEORETICAL IMPORT: IT SAYS THE STUFF OF LANGUAGE HAS A GIVEN DIMENSION (THE FEATURE PARAMETER) AND SOME VALUE (NOT NECESSARILY A BINARY ONE). IMPORTANT QUESTIONS REMAIN ABOUT THE EMPIRICAL SOUNDNESS OF THIS CLAIM. FOR INSTANCE: DO ALL THE LOGICAL POSSIBILITIES THAT FEATURES PERMIT GET INSTANTIATED IN LANGUAGE, AND IF NOT WHY NOT? CAN CERTAIN FEATURES BE DEFINED IN TERMS OF OTHERS, AND IF SO HOW? ARE THERE IMPLICATIONAL RELATIONS AMONG FEATURES, AND IF SO, WHAT CAPTURES THEM?

Again, the assumptions are not innocent. Thus, Epstein et al. (1998) pursue a strong derivational approach in which performance systems access the computation itself, dispensing with levels of representation.

WE COMMENTED ON THIS ABOVE, AND CHOMSKY ACKNOWLEDGES IT. ONE THING TO KEEP IN MIND ABOUT THIS ALTERNATIVE IS THE FOLLOWING. THE EMPIRICAL ARGUMENTS WE HAD AGAINST D-STRUCTURE (KEVIN KEARNEY'S 'TOUGH MOVEMENT' ARGUMENT, AS REPORTED BY LASNIK AND MENTIONED IN CHAPTER 3), DO THEY (OR RATHER, IT) HOLD(S) IN A SYSTEM WITHOUT LEVELS OF REPRESENTATION? IN OTHER WORDS, WAS KEARNEY'S AN ARGUMENT AGAINST THE D-STRUCTURE LEVEL OR AGAINST THE D-STRUCTURE COMPONENT, OR BOTH?

That articulatory and perceptual systems access the same information (PF) is also far from self-evident, corresponding assumptions on the meaning side even less so. And there are many other questions.

THIS IS PRECISELY WHAT WE SAID ABOVE ALSO, AND THE FACT THAT CHOMSKY IS FULLY AWARE OF IT SHOULD BE EMPHASIZED, PARTICULARLY IF YOU WANT TO BUILD A THEORY BASED ON THOSE ASSUMPTIONS.

To say that phonetic features are "instructions" to sensorimotor systems at the interface is not to say that they have the form "move the tongue in such-and-such a way" or "perform such-and-such analysis of signals." Rather, it expresses the hypothesis that the features provide information in the form required for the sensorimotor systems to function in language-independent ways. Similar observations hold on the (far more obscure) meaning side.

THIS SHOULD BE PONDERED SERIOUSLY. THE EQUIVALENT 'DIRECT' INSTRUCTIONS AT LF WOULD BE THINGS LIKE 'DO THIS OR THAT TO REFER' OR 'MOVE HERE OR THERE SO AS TO GET A CERTAIN COMPOSITIONAL MEANING'. RATHER, WE'RE EXPRESSING A GIVEN EMPIRICAL HYPOTHESIS: LF FEATURES PROVIDE INFORMATION IN THE FORM REQUIRED FOR THE CONCEPTUAL/INTENTIONAL SYSTEMS (WITH THE CAVEATS ABOVE) TO FUNCTION IN LANGUAGE-INDEPENDENT WAYS. WE'RE GOING TO HAVE TO EMPIRICALLY DETERMINE WHAT THAT INFORMATION IS (ASSUMING THE HYPOTHESIS IS EVEN RIGHT). FOR INSTANCE, IN THE CASE OF PF, PRESUMABLY CONDITIONS LIKE KAYNE'S LINEARIZATION ARE THERE BECAUSE OF HIGGINBOTHAM'S OLD HYPOTHESIS: THEY ARE DEMANDED BY THE PHYSICS OF SPEECH. THAT'S A RATHER REMOTE INSTRUCTION, BUT A PLAUSIBLE ONE

WHICH DETERMINES SYNTACTIC COMPUTATIONS. PRESUMABLY THOSE EXIST AT LF AS WELL, BUT IT IS MUCH HARDER TO SEE WHAT THEY ARE.

The framework imposes a distinction between (1) linguistic expressions EXP = <PF, LF> that are internal to the mind/brain, and (2) observable events, utterances and actions -- externalization of (mentally-constructed) speech acts.

AGAIN, AN EMPIRICAL DISTINCTION.

No questions arise about the ontological status of the set of expressions {EXP} generated by L;

THOSE WHO WORK WITH 'FORMAL LANGUAGES' OR WITH A LANGUAGE AS A 'SET OF SENTENCES' AND SO ON ARE BEING ADDRESSED HERE. AND NOTE:

its status is somewhat like that of potential visual images or plans for limb motions.

THIS IS ESSENTIALLY THE IDEA OF THE MOTOR SYSTEM I SKETCHED ABOVE, AND IT MAKES LITTLE SENSE TO WORRY ABOUT WHAT ARE THE POSSIBLE OUTPUTS OF THAT SYSTEM, OR SIMILARLY THE VISUAL ONE (IN TERMS OF CONSTRUCTING VISUAL IMAGES). WHAT YOU'RE CONCERNED WITH, REALLY, ARE THE INTERNAL WORKINGS OF THE SYSTEM, AND YOU ONLY USE POSSIBLE OUTPUTS IN DETERMINING THE STRUCTURE OF THE SYSTEM, WHICH YOU DON'T HAVE DIRECT ACCESS TO. OF COURSE, IF YOU COULD PROBE IT DIRECTLY, IT ISN'T EVEN CLEAR THAT YOU WOULD BE VERY CONCERNED WITH OUTPUTS IN ANY SENSE AT ALL.

{Some disagree, regarding the issues as problematic. See, e.g., Carr (1997), and for comment on some related matters, George (1996). On features as "instructions" for vocal gestures, see Halle (1983).}

Finally, I will assume that the Principles-and-Parameters (P&P) approach is in important respects on the right track. Within any version of it, the major problem is to discover the principles and parameters,

THIS SHOULD BE QUALIFIED: WE NEVER REALLY 'DISCOVER' ANYTHING. WE EVALUATE THEORIES THAT POSTULATE HYPOTHESES ABOUT HOW THINGS WORK, AND ON THE BASIS OF OUR EVALUATION WE TENTATIVELY CONCLUDE THAT THIS PRINCIPLE OR THAT PRINCIPLE EXIST. THAT'S WHY IT DOESN'T MAKE MUCH SENSE TO HAVE A 'THEORY OF PARAMETERS' IN THE SENSE OF PREDICTING THE POSSIBLE PARAMETERS. IT ONLY MAKES SENSE TO HAVE SUCH A HYPOTHESIS TO ACTUALLY SEE WHAT'S OUT THERE (SIMILARLY, IT WOULDN'T MAKE SENSE TO HAVE A THEORY TO DISCOVER PRINCIPLES; RATHER, YOU HAVE A THEORY OF WHAT THE PRINCIPLES ARE, PERIOD).

and to show how a particular choice of parameter values and lexicon enters into fixing

THE ACQUISITION TASK, IN PRESENT DAY TERMS. NOTE, THE ISSUE ISN'T TRIVIAL, GIVEN THEORIES OF LANGUAGE CHANGE LIKE KROCH'S OR LIGHTFOOT'S, AND THEIR VARIANTS. YOU WANT TO MAKE SURE THAT THERE IS

ROOM FOR 'FAILURE' IN THE ACQUISITION, AND THEREFORE CHANGE. WHERE THAT FAILURE COMES FROM IS A VERY IMPORTANT QUESTION, AND ANSWERING IT ONE WAY OR THE OTHER HAS TREMENDOUS CONSEQUENCES FOR THE ARCHITECTURE OF PARAMETERS AS WELL AS THE DETAILS OF THE LANGUAGE ACQUISITION DEVICE.

a language L

{L an I-language in the technical sense, here and below. One simplifying assumption is that L is literally deducible from a choice of parameter values and lexicon, so that acquisition is "as if instantaneous." That need not be the case (e.g., in the theory of acquisition proposed by Locke 1997).

AT ISSUE, OF COURSE, IS MATURATION, AND WHETHER THE MAPPING FROM THE INITIAL STATE TO THE ADULT STAGE PROCEEDS THROUGH INTERMEDIATE STAGES, OR INDEED WHETHER PARAMETERS ARE MORE ELABORATE THAN SIMPLE OPEN DIMENSIONS, UNSPECIFIED BY INNATE ENDOWMENT.

It therefore becomes interesting to ask how close to true the assumption is. None of this has anything to do with the existence of a "language acquisition device" (LAD). LAD is just S<0>, under a particular construal, including whatever properties of S<0> may manifest themselves in the course of development. Postulation of LAD is often described as questionable or wrong, but that can hardly be so, at least if language is an identifiable component of human cognitive structure in any respect.};

THAT IS, ONLY THOSE WHO DO NOT BELIEVE IN LANGUAGE AS A COMPONENT OF COGNITION COULD NOT BELIEVE IN A LAD, SINCE PLAINLY LANGUAGE ISN'T INITIALLY THERE AS WE EXPERIENCE IT AS ADULTS, AND THEN IT GETS THERE, BUT IT IS EASY TO SHOW (POVERTY OF THE STIMULUS) THAT WHATEVER WENT ON IN THE ACQUISITION PROCESS ISN'T JUST LEARNING. CHOMSKY IS ALSO SAYING THAT LAD IS NOTHING BUT THE INITIAL STATE OF FL (WITH APPROPRIATE CAVEATS FOR DEVELOPMENT, IF IT INDEED EXISTS IN NON-TRIVIAL WAYS). ONCE YOU HAVE A PRINCIPLES AND PARAMETERS SYSTEM, YOU CAN SOUNDLY SAY THAT: LAD IS WHATEVER YOU HAVE PRIOR TO FIXING THE PARAMETERS, WHICH YOU FIX WITH DATA. THE ONLY TWIST HERE RELATES TO LANGUAGE CHANGE, ALTHOUGH IT IS PERFECTLY POSSIBLE THAT LANGUAGE CHANGE DOES NOT HAVE ANYTHING TO DO WITH LAD PROPER (AND IT INSTEAD INVOLVES POPULATION DYNAMICS, OR SOME SUCH THING, CLEARLY EXTERNAL TO THE LINGUISTIC SYSTEM AT LARGE).

and to proceed beyond, to the study of use, acquisition, pathology, cellular mechanisms, and a wide range of other questions having to do with the place of language in the biological and social worlds.

THIS IS A NICE STATEMENT OF GOALS, RATHER OPEN ENDED AND OPEN MINDED. IT REFLECTS THE MASSIVE PROGRESS IN THOSE FIELDS.

Whatever its ultimate fate, the crystallization of the P&P approach contributed to substantial

progress in several of these areas. The approach also opens some new questions. Concern for descriptive and explanatory adequacy is as old as the study of language. As soon as the two traditional goals were reformulated within modern generative grammar, serious tension arose between them: the search for descriptive adequacy seems to lead to ever greater complexity of rule systems, varying among grammatical constructions and across languages, while the search for explanatory adequacy leads to the conclusion that language structure is largely invariant. It is this tension that has driven the research inquiry of generative grammar from its inception. The P&P framework suggests a way to resolve the tension, thus offering some conception of the form that a genuine theory might take.

INDEED, IT IS ONLY BECAUSE OF THE SUCCESS OF THE P&P MODEL THAT WE'RE EVEN POSING THESE QUESTIONS.

It therefore becomes possible to consider some new questions about FL. In particular, we may ask the question: How well is FL designed?

IN ESSENCE, THAT'S THE KEY TO MINIMALISM. NOTE THAT, UP TO NOW, EVERYTHING WAS GOING AROUND GENERAL CONSIDERATIONS, AND HOW WELL WE'VE DONE IN OUR THEORIZING ABOUT LANGUAGE IN GENERAL AND ITS PLACE WITHIN THE MIND. THE NATURE OF THIS NEW QUESTION, HOWEVER, IS TOTALLY NEW. INDEED, IT IS A SURPRISING QUESTION TO ASK FOR A BIOLOGICAL SYSTEM. IMAGINE WE WERE ASKING 'HOW WELL IS A SPINAL CORD DESIGNED?' WHAT WOULD THAT MEAN? NOBODY DESIGNED A SPINAL CORD! AND INDEED, TO JUDGE FROM BACK-ACHES AND SO FORTH, THE DESIGN ISN'T PARTICULARLY BRILLIANT, WHICH IS WHAT WE EXPECT FROM THE TWISTS AND TURNS OF OUTRAGEOUS EVOLUTION. YET GO ON:

Suppose that a super-engineer were given design specifications for language: Here are the conditions that FL must satisfy; your task is to design a device that satisfies these conditions in some optimal manner (the solution might not be unique). The question is: How close does language come to such optimal design? If the question is real,

WE SHOULD, OF COURSE, WORRY ABOUT THAT. TO THE EXTENT LINGUISTS ARE ASKING THIS QUESTION, WE'RE SEPARATING OURSELVES FROM STANDARD BIOLOGY, AND EVEN STANDARD (COGNITIVE) PSYCHOLOGY, FOR BETTER OR FOR WORSE. MORE ON THIS BELOW.

and subject to inquiry, then the P&P approach might turn out to be an even more radical break from the tradition than it seemed to be.

THAT IS, THE P&P MODEL UNDERSTOOD IN A MINIMALIST LIGHT. NOTHING WITHIN THE PREVIOUS SYSTEM FORCED US TO POSE THE QUESTION WE'RE NOW RAISING. INDEED, SOME VERSIONS OF THAT MODEL WEREN'T PARTICULARLY ATTUNED TO 'GOOD DESIGN', WITH IDLE MOVEMENTS TO SALVAGE A DERIVATION, DELETIONS WHEREVER NECESSARY, DOING THINGS IF NOTHING PREVENTED THEM, AND SO ON. THE SET OF PRINCIPLES THAT ALLOWED FOR THAT WAS QUITE ELEGANT, BUT THE SYSTEM AS A WHOLE WASN'T PARTICULARLY ELEGANT, NOR WAS THERE ANYBODY PARTICULARLY

CONCERNED WITH FINDING A PRETTY SYSTEM. AS A MATTER OF FACT, PEOPLE OFTEN BOASTED OF THE SYSTEM BEING QUIRKY HERE OR THERE, WHICH PROVED ITS BIOLOGICAL ODDITY, THEREFORE ITS NATURALNESS IN THE USUAL WAY.

Not only does it abandon traditional conceptions of "rule of grammar" and "grammatical construction" that were carried over in some form into generative grammar, but it may also set the stage for asking novel questions that have no real counterpart in the earlier study

NOTE THAT CHOMSKY EXPLICITLY THINKS THERE'S A REAL BREAK WITH TRADITION HERE.

of language.{The question of a perfect language, whether designed by God or humans, is of course an old one, but completely distinct. Note further that the question of optimal design has nothing to do with the issue of "best theory" for FL (however intricate and "imperfect" the design of the system).}

THIS IS EXTREMELY IMPORTANT. CHOMSKY IS EXPLICITLY NOT ASKING WHAT IS THE BEST THEORY TO ACCOUNT FOR FL, IN OTHER WORDS, WHAT IS THE METHODOLOGICALLY BEST WAY TO PROCEED. RATHER, HE'S ASKING ABOUT THE DESIGN OF THE SYSTEM, WHICH WE MIGHT IN FACT NEED A RATHER INTRICATE THEORY TO ACCOUNT FOR. IN OTHER WORDS, YOU MIGHT HAVE A MESSY SYSTEM WHICH IS NONETHELESS DESCRIBED BY A THEORY THAT ONLY USES ONE OR TWO ELEGANT PRINCIPLES, OR HAVE, INSTEAD, A VERY ELEGANT SYSTEM WHOSE OPERATION IS HARD TO DESCRIBE OTHER THAN BY WAY OF A MESSY THEORY. THOSE ARE OBVIOUSLY DIFFERENT. THE FIRST WE MAY CALL METHODOLOGICAL MINIMALISM, AND THE SECOND, ONTOLOGICAL MINIMALISM (WITHOUT MAKING ANY CLAIMS ABOUT THE ULTIMATE NATURE OF THIS ONTOLOGY). TO MAKE THESE EXAMPLES CONCRETE, RECALL THE LASNIK AND SAITO SYSTEM ASSUMED IN BARRIERS. THAT SYSTEM WAS METHODOLOGICALLY EXTREMELY ELEGANT: IT ONLY HAD ONE RULE, AFFECT @, WHICH SAID YOU COULD DO ANYTHING TO ANYTHING ANYWHERE ANYTIME IN THE DERIVATION; NO STIPULATIONS HERE. AS A RESULT, THE SYSTEM THAT AROSE WAS HIGHLY INTRICATE: IN CERTAIN DERIVATIONS YOU HAD SEVERAL THINGS MOVING BACK AND FORTH, DELETING, INSERTING, AND SO ON, JUST BECAUSE THE SYSTEM ALLOWED IT. THE PRESENT SYSTEM, AS WE'LL SEE LATER ON, DOESN'T ALLOW YOU TO DO THAT. WHICH MEANS YOU NEED A PRINCIPLE TO PREVENT IT, SAY 'GREED' OR WHATEVER. THE RESULTING THEORY IS MORE COMPLEX: IT HAS THE EXTRA 'GREED' AXIOM. BUT THE SYSTEM IS NOW, AT LEAST IN A SENSE, MORE ELEGANT: ITS BEHAVIORS ARE DRASTICALLY LIMITED IN WAYS THAT ONE COULD CALL 'PRETTY' (AT LEAST THAT'S THE INTENTION--WE'LL SEE WHETHER THE RHETORIC MEETS THE RESULTS).

The "minimalist program" is the attempt to formulate and study such questions. One should bear in mind that it is a program>, not a theory, even less so than the P&P approach. There are minimalist questions, but no minimalist answers, apart from those found in pursuing the program: perhaps that it makes no sense, or that it makes sense but is premature.

KEEP THIS IN MIND. REMEMBER, THE QUESTION IS CRAZY, AS PRESENTLY FORMULATED (MORE ON IT BELOW). CHOMSKY, IN OTHER WORDS, DOESN'T TAKE WHAT HE'S DOING TO BE JUST BUSINESS AS USUAL. THERE IS NO REASON TO THINK IT MAKES NO SENSE TO DO BUSINESS AS USUAL, OR TO THINK THAT IT'S PREMATURE, PARTICULARLY WHEN SO MUCH SUCCESS HAS COME OUT OF IT (IT'S NOT THE ONLY GAME IN TOWN, MIND YOU, BUT THE BEST ONE). BUT WE ARE IN ONE RESPECT PARTIALLY BREAKING FROM THAT TRADITION, FOR BETTER OR FOR WORSE: WE'RE ASKING A DESIGN QUESTION THAT WE HADN'T ASKED BEFORE, AND IT MAY WELL BE JUST A BAD QUESTION. IT IS EXCITING, BUT METHODOLOGICAL SOUNDNESS HERE DEMANDS HEALTHY INSECURITY, EVEN SKEPTICISM (UNLESS EMPIRICAL RESULTS ABOUND).

The program presupposes the common goal of all inquiry into language -- to discover the right theory -- and asks further why language is that way.

THIS IS ANOTHER WAY OF PUTTING IT; WE HAVE NEVER ASKED WHY LANGUAGE IS THAT WAY. IT WOULD HAVE BEEN THE WRONG QUESTION TO ASK IN THE P&P MODEL, WHERE THE METRIC OF SUCCESS WAS A LEARNABLE SYSTEM, A PARSABLE SYSTEM, A CHANGEABLE SYSTEM, AND SO ON; WHY IS IT THAT WAY? WRONG QUESTION: IF IT IS THAT WAY, IT IS LEARNABLE (PARSABLE, ETC.). NOW WE'RE ASKING 'WHY IS IT THAT WAY', AND IN FACT MORE THAN THAT 'WHY IS IT ELEGANT?', OR MORE, STILL: IF YOU TELL ME IT'S NOT ELEGANT, YOU'RE THEORY IS WRONG, SO CHANGE IT TO MAKE THE SYSTEM ELEGANT, AND THEN I ASK WHY IT IS NOW ELEGANT. THAT'S WHAT I'M CALLING CRAZY, AND FURTHERMORE WHAT I'M TRYING TO DEFEND AS A VERY INTERESTING IDEA WHICH I HAPPEN TO BELIEVE MYSELF.

More narrowly, it seeks to discover to what extent minimal conditions of adequacy suffice

DITTO, THAT'S THE CRAZY PART (BIOLOGICALLY, WHY SHOULD THEY BE MINIMAL CONDITIONS, ETC.)

to determine the nature of the right theory.

{It is a misunderstanding to contrast "minimalism and X," where X is some theoretical conception (Optimality Theory, Lexicalism, etc.). X may be pursued with minimalist goals, or not.}

AGAIN, CRUCIAL, AND EQUALLY CRAZY. SO YOU'D BE ASKING WHY IS OPTIMALITY THIS WAY, INDEED, WHY IS IT OPTIMAL? ALL THE ONTOLOGICAL DOUBTS THAT SHOULD INVADE YOU IN THE PREVIOUS INSTANCE TRIVIALLY EXTEND TO THE PRESENT ONE, OR FOR THAT MATTER TO SIMILAR QUESTIONS OUTSIDE THE REALM OF LANGUAGE: 'WHY ARE THE BODY PLANS OF ORGANISMS SO ELEGANT'? (OF COURSE: ARE THEY? ETC.)

Questions of this kind are not often studied, and might not be appropriate at the current level of understanding, which is, after all, still quite thin in a young and rapidly changing approach to the study of a central component of the human brain, perhaps the most complex object in the world,

and not well understood beyond its most elementary properties.

THIS IS AN UNDERSTATEMENT, ALTHOUGH AN OBVIOUS ONE. INDEED, QUESTIONS OF THIS KIND ARE NOT OFTEN STUDIED IN ANY FIELD, CLEARLY NOT IN BIOLOGICAL OR PSYCHOLOGICAL ONES, AND ONLY RARELY IN PHYSICS OR CHEMISTRY. WHY WAS THERE A BIG BANG? WHY IS TIME MOVING FORWARD? WHY DID CHEMICAL COMPOUNDS BECOME POSSIBLE/STABLE? WHY DID LIFE EMERGE 'AGAINST' THE TENDENCY OF THE UNIVERSE TOWARDS DISORDER? THESE ARE NOT OFTEN ASKED, ALTHOUGH THEY ARE SOMETIMES ASKED, IN FIELDS WHICH HAVE A GOOD UNDERSTANDING OF THE COMPONENTS OF THEIR THEORY. THE ONLY GOOD NEWS FOR LINGUISTS IS THAT WE HAVE A VERY GOOD UNDERSTANDING OF THE INTERNAL COMPONENTS OF OUR SYSTEM (WHICH WE DON'T USUALLY GIVE US CREDIT FOR) EVEN IF WE HAVE VERY LITTLE UNDERSTANDING OF THE PHYSICAL SUPPORT OF THIS SYSTEM. IRONICALLY, HARD SCIENTISTS LIVE UNDER THE FICTION THAT IT IS THE OTHER WAY AROUND...

The program is recent, and it is too early to assess it with any confidence. My own tentative judgment has two aspects, one methodological, the other substantive. At the methodological level, the program has a certain heuristic and therapeutic value. It brings to light what might be fundamental problems, where empirical evidence and minimalist expectations conflict. And it encourages us to distinguish genuine explanations from "engineering solutions" -- a term that I do not mean in any disparaging sense. Problems of descriptive and explanatory adequacy are vast and largely obscure. One tries to overcome them somehow, with special assumptions that are often not independently well-motivated, hoping to reformulate the problems in ways that will facilitate further inquiry. Take, say, the study of conditions on extraction of subjects in terms of government and licensing, or attempts to account for the V-second phenomenon or linear ordering in terms of X-bar theory, with its standard stipulations. Various solutions have been proposed that are useful and enlightening, but we can ask whether they are of roughly the order of complexity of the original problem. If so, it would be wrong to conclude that such proposals lack value; on the contrary, they may and often have opened the way to considerable progress. But we can still ask whether they are genuine solutions. Or consider the target of V-raising. Evidence has accumulated that it can be to a position higher than tense but lower than C, differentiating languages by the position of a functional category ", on current assumptions.

{See Belletti (1990), and much subsequent work.}

But there is reason to doubt that such " can exist; or to put it differently, if it does then departures are needed from what appears to be the simplest and most principled form of phrase structure theory. Again questions arise as to whether there is some better way to conceive the matter. The minimalist program helps to focus attention on such issues, and perhaps to address them by showing that elimination of descriptive technology yields empirical results that are as good, possibly even better, than before.

THIS IS THE PART OF THE SYSTEM THAT IS JUST BUSINESS AS USUAL: YOU WANT TO SEE HOW ADEQUATE YOUR THEORETICAL PARAPHERNALIA IS, AND YOU USE WHATEVER TOOL YOU CAN GET TO DECIDE IN THIS. THERE IS A TWIST, THOUGH: THE TOOL IN THIS INSTANCE IS THE CRAZY LITTLE IDEA THAT IF YOU DO THIS

OR THAT, THEN YOU DEPART FROM THE SIMPLEST SUCH-AND-SUCH. IF THAT IS MEANT IN MERE OCKHAM'S RAZOR'S TERMS, THEN IT MAKES (STANDARD) SENSE, PARTICULARLY IF I CAN ACCOUNT FOR THE SAME FACTS WITH LESS THEORY. BUT THE CLAIM IS MADE IN STRONGER TERMS:

The substantive thesis is that language design may really be optimal in some respects, approaching a "perfect solution" to minimal design specifications.

THAT IS, 'IF YOU DO THAT, THE RESULTING SYSTEM WOULD BE SUB-OPTIMAL (EVEN IF THE THEORY BEHIND IT IS OPTIMAL); SO DON'T DO THAT.' THAT'S THE CRAZY, ONTOLOGICAL (OR SUBSTANTIVE) PART:

The conclusion would be surprising, hence interesting if true.

BOTH OF THESE SHOULD BE OBVIOUS.

<2. Design Specifications>

Proceeding along the course just outlined, we face two questions: (1) What is "good design"? (2) What are the minimal design specifications for FL? The former we may put to the side, not because it is unimportant, but because questions of this nature are common to all rational inquiry. The appropriate place to seek answers is in the hard sciences, where understanding is far deeper and intuitions far more firmly grounded.

THIS IS ALL OBVIOUS, BUT THERE'S ONE THING TO KEEP IN MIND. UNLESS WE TAKE THE QUESTION SOMEWHAT SERIOUSLY, WE MAY BE CONFUSING OUR FAVORITE PET THEORY WITH 'GOOD DESIGN'. THERE ARE NO PANACEAS HERE, BUT AT LEAST SOME MINIMAL SENSITIVITY TOWARDS WHAT COUNTS AS GOOD DESIGN IS APPROPRIATE.

Caution is in order in appealing to such considerations. Given some empirically-supported conclusion, it is often possible to construct plausible conceptual grounds for it; and for alternatives.

INDEED, THAT MIGHT BE A GOOD STRATEGY. IF YOU CLAIM X IS A GOOD DESIGN, HOW ABOUT ITS ALTERNATIVE(S). ARE THEY ALSO GOOD DESIGN? OF COURSE, NO HARD ANSWERS HERE EITHER--PERHAPS THERE ARE SEVERAL GOOD DESIGNS. BUT AT ANY RATE, FAMILIAR QUESTIONS.

To clarify the problem of design specifications, let us invent a evolutionary fable, keeping it highly simplified.

{Complications can readily be added. Little is known about evolution of higher mental faculties, and it is not clear how much can be learned within the limits of contemporary understanding; for a skeptical appraisal, see Lewontin (1990, 1998), and for critical analysis of recent efforts, Berwick (1997), Jenkins op. cit>., Orr (1997).}

Imagine some primate with the human mental architecture and sensorimotor apparatus in place, but no language organ. It had our modes of perceptual organization, our propositional attitudes (beliefs, desires, hopes, fears,...) insofar as these are not mediated by language, perhaps a

"language of thought" in Jerry Fodor's sense, but no way to express its thoughts by means of linguistic expressions, so that they remain largely inaccessible to it, and to others.

RECALL THAT, IN PRINCIPLE, NONE OF THIS IS PARTICULARLY INSANE OR EVEN IMPLAUSIBLE, SINCE SUB-COMPONENTS OF THE LANGUAGE FACULTY MIGHT HAVE BEEN THERE FOR TOTALLY DIFFERENT REASONS.

Suppose some event reorganizes the brain in such a way as, in effect, to insert FL.

THIS MIGHT BE SOMETHING AS TRIVIAL AS, SAY, THE ABILITY TO LINEARIZE IN KAYNE'S SENSE. WITHOUT THAT, YOU OBVIOUSLY HAVE NO FL IN ANY USEFUL SENSE. WHY IT (WHATEVER IT IS) GOT IN PLACE, GOD ONLY KNOWS; CERTAINLY, 'COMMUNICATIVE PRESSURES' WOULD HAVE NOTHING TO DO WITH IT, FOR REASONS ALREADY MENTIONED. TO ELABORATE ON THIS SLIGHTLY, THE ABILITY, FOR INSTANCE, TO RHYTHMICALLY ASSEMBLE SOUND ITEMS MIGHT HAVE HAD TREMENDOUS EVOLUTIONARY ADVANTAGES IN TERMS OF PARENT/CHILD RELATIONS (LULLABIES...), RELATIONS AMONG THE SEXES (COURTING RITUALS...), SOCIAL BONDING AMONG ADULTS (WORKING OR HUNTING RITUALS...), CARING FOR THE SICK AND ELDERLY (SOLACE TUNES), OR EVEN BURIAL AND SUPERNATURAL CELEBRATION. (LISTEN TO SO CALLED PRIMITIVE MUSIC AND YOU'LL GET ALL THAT, AND INCIDENTALLY, THAT KIND OF MUSIC OFTEN HAS NO WORDS, OR ONLY SCATTERED WORDS WITHOUT OBVIOUS REFERENTS.) THOSE ALONE MIGHT JUSTIFY ADAPTING A RHYTHMIC SYSTEM, WHICH IN THE TERMS OF JACKENDOFF AND LERDAHL 1984 MAPS HIERARCHICAL STRUCTURES TO LINEAR SEQUENCES (IS AKIN TO KAYNE'S LCA). THE POINT IS, YOU MAY HAVE GOTTEN THE LCA FOR TOTALLY DIFFERENT REASONS; ONCE THE LCA IS THERE, THOUGH, IT MAY HAVE 'LIBERATED' LANGUAGE USE AS WE NOW KNOW AND LOVE IT.

To be usable, the new organ has to meet certain "legibility conditions."

{These are called "bare output conditions" in MP>, "output" because they are conditions on interface levels, hence "outputs" on a derivational approach; "bare" to distinguish them from filters, ranked constraints, and other devices that are part of the computational system itself.}

'LEGIBILITY' IS A BETTER TERM THAN 'INTERPRETABILITY', WHICH CAN GET CONFUSED WITH 'INTELLIGIBILITY'.

Other systems of the mind/brain have to be able to access expressions generated by states of FL ((I-)languages), to "read" them and use them as "instructions" for thought and action. We can try to formulate clearly -- and if possible answer -- the question of how good a solution FL is to the legibility conditions, and these alone. That is essentially the topic of the minimalist program.

COULDN'T BE MORE CLEARLY EXPRESSED. THIS NARROWS YOUR FIELD OF RESEARCH, IF YOU TAKE IT SERIOUSLY. IN ESSENCE, YOU MUST ASSUME AS PART OF THE PROGRAM WHAT THE LEGIBILITY CONDITIONS ARE, AND OF COURSE, SOME ANSWER TO THE QUESTION OF HOW WELL THE SYSTEM

ADDRESSES THOSE LEGIBILITY CONDITIONS. IN OTHER WORDS, WHAT YOU'RE LOOKING FOR ARE NATURAL INTERACTIONS WITH OUTSIDE SYSTEMS, AND INTERNAL COHERENCE (INDEED ELEGANCE) WITHIN THE SYSTEM THAT MATCHES THOSE OUTSIDE SYSTEMS AT THE LEGIBILITY POINT(S).

We have assumed two external systems: sensorimotor systems and systems of thought, each with its own characteristics independently of FL. The former can only use information presented in a specific form: with temporal order, prosodic and syllable structure, certain phonetic properties and relations. The systems of thought require information about units they can interpret and the relations among them: certain arrays of semantic features, event and quantificational structure, and so on. Insofar as we can discover the properties of these external systems (an empirical problem, however difficult),

THAT'S THE 'NATURAL INTERACTIONS' BIT, NOT TO BE ANSWERED BY AUTHORITY ARGUMENTS...

we can ask how well

AGAIN, WITHIN CERTAIN UNCLEAR PARAMETERS, FAMILIAR TO THE HARD SCIENCES

the language organ satisfies the design specifications they impose, providing legible representations at the interface levels. That is the minimal condition FL must satisfy to be usable at all.

{For significance, we might assume further that there is no (nonarbitrary) bound on the number of legible expressions.

I HONESTLY DON'T UNDERSTAND WHAT 'FOR SIGNIFICANCE' MEANS, AND WON'T VENTURE SPECULATIONS. THE LACK OF BOUND, OF COURSE, IS A FACT, AND A DEEP ONE. PERSONALLY, I'D TRY TO SAY THAT THERE'S SOMETHING DEEPLY ELEGANT ABOUT THAT, IN FACT A PROPERTY THAT FL SHARES WITH (ONLY?) FRACTAL SYSTEMS IN NATURE, WHICH HAPPEN TO BE ELEGANT IN WAYS THAT FEW OTHER SYSTEMS CAN BE SAID TO BE.

Note that FL satisfying this minimal condition might -- and the real system in fact does -- permit generation of expressions that are unusable (structure of memory, garden path, etc.).}.

THAT'S IMPORTANT, SINCE IT SHOWS A MISMATCH BETWEEN WHAT'S KNOWN AND WHAT'S USABLE, RELATING TO SOMETHING MENTIONED ABOVE REGARDING THE NATURE OF THE COMPETENCE/PERFORMANCE SPLIT.

To introduce some terminology of MP>, We say that a computation of an expression EXP converges at an interface level IL> if EXP is legible at IL,

BEFORE WE USED TO SAY 'INTERPRETABLE', AN UNFORTUNATE TERM.

consisting solely of elements that provide instructions to the external systems at IL and arranged so that these systems can make use of them; otherwise it crashes> at IL.

IT'S A PITY WE HAVEN'T USED THE TERM 'DIVERGES', WHICH WOULD EMPHASIZE THE FACT THAT THE REPRESENTATION IS WHATEVER IT IS, AND JUST DOES NOT CONVERGE AS A LEGIBLE LF OR PF.

The computation converges> if it converges at all interfaces. Call the expression EXP so formed convergent> as well. As in MP>, we keep here to a restricted version of the concept of convergence, setting aside the matter of legible arrangement (which raises all sorts of complex issues), and tentatively assuming it to be irrelevant -- no slight simplification.

I DON'T KNOW WHAT A 'LEGIBLE ARRANGEMENT' WOULD BE, WHETHER IT'S PART OF A LEVEL OR WHAT. I HAVEN'T SEEN THE EXPRESSION USED BEFORE, ALTHOUGH I SUSPECT IT DOES REFER TO AN ARTICULATED SET OF STRUCTURES. IF SO, THAT NOTION WOULD BE NEEDED IN A SYSTEM THAT DOES NOT WORK WITH ARTICULATED LEVELS OF REPRESENTATIONS, AND WHICH THUS MUST DECIDE ON THE CONVERGENCE OF STRUCTURED SETS OF SYMBOLS IN TERMS OF A SMALLER DOMAIN OF LEGIBILITY.

Certain features of lexical items are interpretable>, that is, legible to the external systems at the interface; others are uninterpretable>.

BACK TO THE UNFORTUNATE TERM. I DON'T SEE WHY WE COULDN'T HAVE CALLED THE LEGIBLE FEATURES 'LEGIBLE', ETC.

We assume, then, that if an expression contains only features interpretable at IL, it converges at IL.

{Interpretability is not to be confused with intelligibility. A convergent expression may be complete gibberish, or unusable by performance systems for various reasons. See preceding note. And performance systems typically assign interpretation to nonconvergent expressions.}

AGAIN, THIS REVOLVES AROUND THE SAME TERMINOLOGICAL POINT. AT ANY RATE, THE ONLY SIGNIFICANT THEORETICAL POINTS ARE LEGIBILITY AND INTELLIGIBILITY, AS WELL AS THE FACT THAT PERFORMANCE SYSTEMS CAN IMPOSE INTERPRETATIONS (ALL SORTS OF THEM) ON ILLEGIBLE CONSTRUCTIONS, AS SEEMS NATURAL IF THEY ARE INDEPENDENT, AND AGAIN KEEPING IN MIND THAT THIS WILL NOT MEAN THAT THE RELEVANT, INTERPRETED STRUCTURES ARE, IN ANY SENSE, 'GRAMMATICAL'. AFTER ALL, WE CAN INTERPRET TARZAN SPEAK, EVEN IF WHAT HE SPEAKS ISN'T REALLY ENGLISH.

The property converges at IL> may hold of an expression formed in the course of a derivation that then proceeds on to IL. If, say, particles or adverbs have only LF-interpretable features, then they converge at LF when extracted from the lexicon and at every subsequent stage of derivation to LF.

THIS IS A SOMEWHAT ODD DEFINITION OF CONVERGENCE, IMPOSED BY THE FACT THAT THE SYSTEM CHOMSKY PURSUES HAS LEVELS OF REPRESENTATION. THE IDEA IS THIS: IF X IS CONVERGENT AT LF, SAY, YOU DON'T HAVE TO WAIT UNTIL LF TO DECIDE ON X'S CONVERGENCE. CONVERGENCE THEN MEANS

SOMETHING LIKE: 'YOU'RE THE TYPE OF OBJECT THAT WILL BE OKAY AT LF.' IN A SYSTEM WITHOUT LEVELS THE ISSUE DOESN'T OBVIOUSLY ARISE, BUT YOU WILL HAVE TO DECIDE ON 'LEGIBLE ARRANGEMENTS' (ASSUMING THAT'S THE TERM CHOMSKY HAS IN MIND).

Similarly, an embedded clause may converge, for example, the bracketed subpart of "John thinks [it is raining]." The phrase "converge at an interface" should not mislead: convergence is an internal property of an expression, detectable by inspection.

{Convergence is defined in terms of properties of the external systems; the concept is clear insofar as these properties are clear. Many questions arise about its role in interpretation of deviance and in economy conditions, specifically, does crash "free up" alternative derivations, as assumed in MP>

THIS WAS A SIGNIFICANT PROPERTY OF MP: THE SET OF DERIVATIONS TO ENTER THE OPTIMALITY RACE DOES NOT INCLUDE NON-CONVERGENT ONES.

but not as these notions are elaborated elsewhere (e.g., Collins 1997, and in what follows).}

WHAT FOLLOWS IS A VERY DIFFICULT PASSAGE, WHICH I DON'T REALLY CLAIM TO FULLY UNDERSTAND. LET'S TRY:

Suppose that in state L, FL generates expressions EXP = <PF, LF>. Then L determines sound-meaning associations: the sounds and meanings determined by PF and LF, respectively, are associated in EXP. These are matters of fact that lie well beyond legibility conditions.

AS I UNDERSTAND THIS, THE ISSUE HERE IS ONE OF INTELLIGIBILITY, NOT LEGIBILITY. WHAT MAKES A SOUND-MEANING PAIR INTELLIGIBLE IS A MATTER OF FACT, BUT IT INCLUDES ISSUES THAT GO BEYOND WHAT THE GRAMMAR GENERATES. SO FOR EXAMPLE, THE SENTENCE who does Mary love John? IS PRESUMABLY GRAMMATICAL (NO OBVIOUS VIOLATION OF ANYTHING; NOTE, IF YOU'RE TEMPTED TO SAY THAT IT VIOLATES CASE CHECKING BECAUSE OF WHO, THIS ISN'T OBVIOUS THEORETICALLY--IT IS NOT CLEARLY THE FORMATIVE who, SPECIFICALLY, THAT HAS TO HAVE ITS CASE FEATURES CHECKED (THUS YOU COULD INSERT IT IN THE NUMERATION WITHOUT A CASE FEATURE, IF YOU WANT TO BE TECHNICAL)--AND IN ANY CASE WOULD POSE AN EMPIRICAL PROBLEM WITH who do you wonder whether Mary loves him (WHERE who CLEARLY DOES NOT CHECK CASE). SO IF YOU WANT TO, CONSIDER who do you wonder whether Mary loves John. I DO NOT KNOW OF ANYTHING THAT THE LATTER SENTENCE VIOLATES, AND HENCE I WOULD BE FORCED TO SAY THAT IT CONVERGES, THUS IS LEGIBLE. IS IT INTELLIGIBLE, THOUGH? PRESUMABLY NOT. CONVERSELY, YOU MAY HAVE ILLEGIBLE SENTENCES THAT ARE PERFECTLY INTELLIGIBLE; THUS TARZAN'S 'Tarzan love Jane' DOESN'T CONVERGE IN TERMS OF CASE, YET IS PERFECTLY INTELLIGIBLE TO ANYONE.

{Also, well beyond the minimal sound-meaning connection given by initial assembly of features in the lexicon, an inescapable step in fixing a language for familiar reasons.}

I SUPPOSE CHOMSKY'S TALKING ABOUT THE ARBITRARY PAIRINGS OF SOUND

AND MEANING THAT CONFORM WORDS. INTELLIGIBILITY CONDITIONS PROBABLY GO BEYOND THAT. INCIDENTALLY, WHAT ARE THE FAMILIAR REASONS WHY LANGUAGE IS FIXED THAT WAY, I HONESTLY DON'T KNOW. I SUPPOSE CHOMSKY IS REFERRING HERE TO SAUSSUREAN ARBITRARINESS, BUT IT WOULD SEEM TO ME THAT IT IS A HIGHLY MYSTERIOUS PROPERTY OF LANGUAGE THAT ITS FEATURES SHOULD ARRANGE THEMSELVES INTO ARBITRARY WORDS.

Take such standard examples as (1):

(1)

(i) John is impossible to forgive

(ii) John is impossible to be forgiven

Suppose L assigns to both a semantic representation LF on the model of "John is unlikely to forgive." Then the generated expressions converge but with the wrong pairings.

WHAT'S BEING ASKED OF YOU HERE, I THINK, IS HIGHLY HYPOTHETICAL AND DIFFICULT TO ENTERTAIN. FORGET WHAT YOU KNOW ABOUT THE STRUCTURE OF THE SENTENCES IN (1), AND BE NAIVE ABOUT IT (IN OTHER WORDS, LET'S TRY TO MAKE THEM INTELLIGIBLE IN ANY WAY WE CAN). YOU ASSUME THAT THE SEMANTICS OF THOSE SENTENCES ARE BASED ON THE STRUCTURE OF 'JOHN IS UNLIKELY TO FORGIVE' (IMAGINE YOU'RE NOT A NATIVE ENGLISH SPEAKER, WHO MAKES THIS AS A REASONABLE GUESS). UNDER THE ASSUMPTION THAT 'JOHN IS UNLIKELY TO FORGIVE' IS GOOD, YOUR LFS FOR THE SENTENCES IN (1) WILL CONVERGE, OBVIOUSLY. HOWEVER, THEY ARE INAPPROPRIATELY PAIRED WITH THE ACTUAL SENTENCES WE HAVE.

Suppose L assigns to both a representation LF' corresponding (closely enough) to that of "it impossible to forgive John."

I SUPPOSE THERE'S A TYPO HERE, AND WE MEAN 'IT IS IMPOSSIBLE TO FORGIVE JOHN'

Then the association is right for EXP<1> = <PF<1>, LF'> and wrong for EXP<2> = <PF<2>, LF'>,

NOTE THE LOWER VERB IS PASSIVE IN THE SECOND INSTANCE, HENCE A VERY DIFFERENT STRUCTURE IS INVOLVED (SAY, NO SMALL v, AN AUXILIARY, PASSIVE MORPHOLOGY AND WHATEVER IT MEANS).

though both converge.

SAME AS ABOVE, EXCEPT YOU'RE USING A DIFFERENT HYPOTHETICAL LF HERE.

The assignment is wrong because it does not indicate the deviance of EXP<2>, a crucial property distinguishing it from EXP<1>.

AFTER ALL, THE SECOND INSTANCE IS BAD, AND THE GRAMMAR HAS TO RECOGNIZE THAT, AND CANNOT SKIP IT BY ASSIGNING IT A CONVERGENT LF.

The conclusion holds even if the language user in state L assigns LF' to PF<2> by some interpretive mechanism, using L but presumably going beyond.

AS I UNDERSTAND THE 'GOING BEYOND' PART, THE IDEA IS THIS: YOU CANNOT REACH (ii) IN AN IMPROPER WAY, AND THEN BY SOME ASSOCIATION THAT 'GOES BEYOND' IN SOME SENSE, YOU ADAPT TO A CONVERGENT, SIMILAR, STRUCTURE. YOU WANT TO PREVENT THAT.

Suppose all "best ways" to satisfy legibility conditions yield incorrect associations. Then departure from optimal design is required.

I SUPPOSE THE NEXT SENTENCE IS MEANT AS: 'ON THE OTHER HAND,...'

If there are some that consistently yield the right sound-meaning relations, then we have reason to believe that language design is optimal in non-trivial respects.

SO AS I READ THIS, THE POINT IS SIMPLE. IF THE ONLY WAY LEGIBILITY CONDITIONS ARE MET IS BY TAKING A NON-OPTIMAL PATH, THEN TOO BAD FOR THE IDEA THAT LANGUAGE IS OPTIMAL; BUT IF WHAT YOU FIND OUT IS THE OPPOSITE, AND INDEED LEGIBILITY CONDITIONS CAN BE MET IN SOME OPTIMAL WAY IN A CONSISTENT FASHION, THEN THE IDEA THAT LANGUAGE IS OPTIMAL WILL BE ON THE RIGHT TRACK.

Suppose that FL satisfying legibility conditions in an optimal way satisfies all other empirical conditions too: acquisition, processing, neurology, language change,... Then the language organ is a perfect solution to minimal design specifications.

NOTE, THIS RAISES THE STAKES (FROM LEGIBILITY TO EVERYTHING ELSE). OF COURSE, THAT'S THE CRAZY IDEA BEING ENTERTAINED HERE (AND THE MORE CONNECTIONS YOU MAKE TO ACQUISITION, PROCESSING, AND SO ON, THE CRAZIER IT SOUNDS). AS I UNDERSTAND IT, THE POINT HERE IS EVEN MORE RADICAL: PERHAPS *ALL* THERE IS TO LANGUAGE IS LEGIBILITY CONDITIONS MET IN OPTIMAL TERMS; IF THAT ALSO ACCOUNTS FOR ACQUISITION, PROCESSING, ETC., THEN YOU'D HAVE A VERY ROBUST MINIMALIST THESIS.

That is, a system that satisfies a very narrow subset of empirical conditions in an optimal way -- those it must satisfy to be usable at all

DON'T GET CONFUSED HERE: THOSE ARE THE GRAMMATICAL CONDITIONS WITHIN CURRENT ASSUMPTIONS.

-- turns out to satisfy all empirical conditions. Whatever is learned about other matters will not change the conclusions about FL.

OF COURSE, IF THIS WERE TRUE, PARSING AND PSYCHOLINGUISTIC RESEARCH IN GENERAL, NEUROLINGUISTIC STUFF, AND SO ON, WOULD BE ONLY INDIRECTLY RELATED TO LINGUISTICS PROPERLY. AT LEAST, SO THE RHETORIC

GOES. BUT BE CAREFUL HERE AGAIN. THAT KIND OF RESEARCH, FOR WHAT GOES ON AS WE SPEAK (ON LINE) MAY BE TOO REMOTE TO DETERMINE THE PROGRAM ANYWAY--TOO MANY COMPLICATING FACTORS. ON THE OTHER HAND, SUPPOSE YOU'RE ASKING THE ULTIMATE-ULTIMATE QUESTION (EVEN MORE INSANE THAN THE OTHERS) OF WHY THE LANGUAGE FACULTY TURNED OUT TO BE THE WAY IT DID--SAY, AN OPTIMAL REALIZATION OF LEGIBILITY CONDITIONS. THEN, AGAIN, NEUROLINGUISTIC AND PSYCHOLINGUISTIC STUFF MAY AGAIN BE RELEVANT. FOR EXAMPLE, JIM REGGIA SPECULATES WITH THE POSSIBILITY THAT LANGUAGE STRUCTURE WOULD HAVE TO HAVE EVOLVED IN STRUCTURALLY OPTIMAL WAYS IN ORDER FOR THE RELEVANT NEURAL NETWORKS TO HAVE EMERGED AS OPTIMAL (AND UNDER CERTAIN ASSUMPTIONS, EVEN POSSIBLE) SYSTEMS. THAT MIGHT BE RIGHT OR WRONG, AS AN EMPIRICAL CLAIM, BUT IT POSES THE QUESTION IN SUCH A WAY THAT NEUROLINGUISTICS, PSYCHOLINGUISTICS, COMPUTATIONAL LINGUISTICS, AND SO FORTH, IS AGAIN VERY RELEVANT, AT THIS LEVEL OF ABSTRACTION. THAT HAS TO DO WITH LANGUAGE EVOLUTION, BUT SIMILAR QUESTIONS CAN BE POSED ABOUT LANGUAGE DEVELOPMENT AND, TO START WITH, WHETHER IT EXISTS IN NON-TRIVIAL WAYS. IF IT DID, ONE WOULD THEN HAVE TO WORRY ABOUT WHAT IN THE SEQUENCE OF ACQUISITION CORRESPONDS TO WHAT IN THE MATURING INDIVIDUAL. THE POINT IS, THESE TWO QUESTIONS (EVOLUTION AND DEVELOPMENT) CAN BE SEEN AS PRIOR TO THE QUESTION OF OPTIMALLY MET LEGIBILITY CONDITIONS THAT WORRIES CHOMSKY--WHICH DOESN'T MEAN, INCIDENTALLY, THAT YOU HAVE TO ANSWER THOSE QUESTIONS TO DEDICATE YOURSELF TO THE USUAL ONES. NEEDLESS TO SAY, VERY LITTLE, WITHIN PRESENT DAY DISCIPLINES, HAS CONTRIBUTED TO CLARIFYING ANY OF THESE ISSUES. SO BACK TO LINGUISTICS FOR NOW.

That would be a strange and surprising result, therefore interesting to whatever extent it might be true. The minimalist program explores the possibility that language approaches "good design" in this sense.

THE CAVEAT 'APPROACHES' HERE IS SIGNIFICANT... REMEMBER, OURS IS AN EMPIRICAL EXERCISE ON LIMITS. IF IN THE LIMIT THE SYSTEM WOULD BE PERFECT, BUT IT IS NOT, ACTUALLY, THIS WILL SIMPLY MEAN THAT NOTHING OUT THERE IS COMPLETELY PERFECT. AGAIN, PERFECTLY COMPATIBLE WITH THE LOGIC OF EVOLUTION, AND NOT NECESSARILY A BAD RESULT (THE IMPERFECT, TINY IMBALANCE OF PROTONS AND ANTIPROTONS LEAD TO THE PRESENT UNIVERSE...)

The strongest minimalist thesis would be that:

(2) Language is an optimal solution to legibility conditions

Insofar as the thesis is true, information about other matters (sound-meaning connections, neurophysiology, etc.) may be helpful in practice -- even indispensable -- for discovering the nature of FL and its states. But it is irrelevant in principle.

THIS IS WHAT I JUST SAID ABOVE. SYNTAX, IF THIS IS TRUE, CARRIES THE LEAD

IN A RADICAL WAY. NONETHELESS, DON'T FORGET THE OTHER ISSUES CONCERNING *WHY* THE SYSTEM SHOULD HAVE MET LEGIBILITY CONDITIONS IN AN OPTIMAL WAY, ETC. AT THOSE LEVELS YOU MAY AGAIN NEED NEUROPHYSIOLOGY OR WHATEVER, IN SERIOUS WAYS. (INCIDENTALLY, THE CONVERSE IS A MORE IMPORTANT POINT TO MAKE FOR NEUROPHYSIOLOGISTS: IT'S CRAZY TO TRY TO ANSWER SOME OF THE QUESTIONS THEY ARE TRYING TO ANSWER WITHOUT TAKING SERIOUSLY WHAT LINGUISTS HAVE TO SAY ABOUT HOW THE SYSTEM WORKS; WHEN I WAS TAKING ABOUT AN ENLIGHTENED NEUROPHYSIOLOGY THAT WOULD HELP DECIDE ON THE DEEP ISSUES, I HAD IN MIND SOMETHING WHICH, SO FAR AS I KNOW, PRACTICALLY DOESN'T EXIST--POEPPLE AND A COUPLE OF FRIENDS...)

The tasks of biology of language remain as before, but become even more intriguing and difficult, because a new problem arises: how did the structure of the brain and the course of evolution happen to yield the outcome (2)?

FROM THE PERSPECTIVE OF ONTOLOGICAL MINIMALISM, THAT'S THE DEEPEST QUESTION, PARTICULARLY GIVEN THE NEO-DARWINIAN 'SYNTHESIS'. AT ANY RATE, THAT'S WHERE YOU AGAIN NEED TO HELP OF SERIOUS BIOLOGISTS, WORKING IN TEAMS WITH LINGUISTS, ETC.

The internalist study of language -- syntax in the broad sense -- becomes much harder, hence more interesting and significant, because a standard of explanation is set that is very difficult to meet: descriptive machinery must satisfy stringent conditions, imposed by (2). Issues relating to the interface become of central concern. The problem of discovering whether, and if so how, considerations of economy enter into language design also gains new prominence, along with questions about their role in language acquisition. In general, all questions become harder, hence more interesting and significant -- insofar as there is some truth to the strong thesis. In these terms we might also be able to devise an interpretation for a thesis about language and psychology that seems to make little sense, but that has been so widely held that one might suspect that some significant intuition may lie behind it.

WHAT YOU'RE ABOUT TO SEE IS A VERY FUNNY ARGUMENT, WHICH TURNS AROUND TRADITIONAL CRITICISMS FROM PSYCHOLOGISTS INTO AN 'IN YOUR FACE' RESULT, SADLY ONLY HYPOTHETICAL...

The thesis is that linguists are to study "linguistic evidence" and "linguistic intuitions," but the results of their work, however revealing and far-reaching, do not bear on "reality," sometimes called "psychological reality." Other kinds of evidence are required to find out about reality. To take a classic example, Sapir provided rich "linguistic evidence" for phonological analyses he proposed, and went on to adduce much weaker "psychological evidence" to demonstrate their "psychological reality"; this was considered an audacious and controversial move, mainly on grounds that even the psychological evidence doesn't bear on reality. In the recent period, such ideas have appeared often in critical discussion of the program of generative grammar. Similarly, it is sometimes held that conclusions based on linguistic evidence must be confirmed by "converging" evidence from other sources, though conclusions based on these sources stand on their own. Another variant, thankfully put to rest after too many years, is that it is the task of psychologists to test (verify, refute) the theories of linguists, which are based on "linguistic

evidence," not to contribute directly to these theories. The linguistic evidence is generally understood to consist of informant judgments about sound and meaning and their relations.

{I put aside a variant that restricts "linguistic evidence" to identification of "well-formed" ("grammatical") expressions, so the linguist then faces the alleged problem of selecting among grammars that are extensionally equivalent over these objects. Such demands inherit (and, by the radical restriction of evidence, amplify) the incoherence of the other approaches, adding the further difficulty of deciding what this property might be, for natural language.}

The proposal is odd, as has often been pointed out: evidence does not come with a mark saying "I do or do not bear on reality." Judgments about (1i,ii) have no different status than other kinds of evidence with regard to the nature of FL, as part of the (physical) world. As in the study of vision and other domains, these judgments are discovered by experiment, typically informal in this case, though they can be done as carefully as is necessary to advance understanding. Uncontroversially, one seeks the widest possible range of relevant evidence, converging or conflicting. But there is no principled asymmetry between categories of evidence in this regard.

UP TO HERE, THE TRADITIONAL (SOUND) DEFENSE AGAINST THOSE ATTACKS FROM EMPIRICISM, BEHAVIORISM, AND SO ON.

In terms of the preceding discussion, we might replace these proposals by a substantive (but extraordinarily strong) empirical hypothesis, namely, the thesis (2): an optimal solution to legibility conditions satisfies all other empirical tests as well. The reformulated thesis replaces the obscure notion of "linguistic evidence" by the meaningful notion: satisfaction of interface conditions. Relevant evidence is very limited. The thesis is even stranger than the requirement that "linguistic evidence" suffices to determine grammars, but it has empirical content. One might, perhaps, suggest it as a kind of rational reconstruction of dubious ideas about linguistic evidence and psychological reality that have appeared in one or another form.

THIS IS THE BRAVE NEW APPROACH. SO FINE: THIS ISN'T PSYCHOLOGY OR ANYTHING OF THE SORT. IT'S JUST 'MEETING LEGIBILITY CONDITIONS'. IT'S A NATURAL SYSTEM WITH THAT AS ITS CENTRAL PROPERTY.

Suppose we understood external systems well enough to have clear ideas about the legibility conditions they impose. Then the task at hand would be fairly straightforward at least to formulate: construct an optimal device to satisfy just these conditions, and see how well it satisfies other empirical conditions. If all such efforts fail, then add "imperfections" as required.

THAT'S THE IDEALIZED PROGRAM, OF COURSE, WHERE AS YOU SEE PSYCHOLOGICAL (OR FOR THAT MATTER BIOLOGICAL, PHILOSOPHICAL, MATHEMATICAL ...) EVIDENCE (IN ANY TRADITIONAL SENSE) IS ENTIRELY BESIDES THE POINT, IN PRINCIPLE.

But life is never that simple. The external systems are not well understood. Progress in understanding them goes hand-in-hand with progress in discovering the language systems that interact with them. So the task is simultaneously to set the conditions of the problem and to try to satisfy them, with the conditions changing as we learn more about how to do so.

THANK GOD, OR THE PROGRAM WOULD BE SOLVED IN LESS TIME THAN THE

GENOME PROJECT... AND WE'D BE OUT OF WORK.

That is not surprising. It is much what we expect when trying to understand some complex system.

I THINK 'COMPLEX' HERE IS MEANT SERIOUSLY. MORE ON THIS BELOW.

We proceed with tentative proposals that seem reasonably firm, expecting the ground to shift as more is learned.

<3. Architecture>

We are taking L to be the recursive definition of a set of expressions EXP = <PF, LF>. We can now raise a question -- at least, an apparent question -- about the interpretation of the recursive definition. One might construe L as a step-by-step procedure for constructing EXPs, suggesting that this is how things work as a real property of the brain, not temporally but as part of its structural design.

THIS LITTLE SIDE REMARK IS IMPORTANT: DERIVATIONAL DOESN'T HAVE TO MEAN TEMPORAL. CONSIDER, FOR INSTANCE, A FAMILY TREE, YOU, YOUR PARENTS, AND SO ON. YOU CAN EASILY COME UP WITH A DERIVATIONAL SYSTEM THAT DESCRIBES THAT TREE, STARTING WITH YOU, AND GOING ALL THE WAY DOWN TO SOME SET OF HOMINIDS DOWN IN AFRICA. BUT IT IS EASY TO SEE THAT THIS DERIVATION DOESN'T ADVANCE TEMPORALLY.

Assumptions of this nature constitute a derivational approach> to L. The strong> derivational approach dispenses with the expression altogether, assuming that information is provided to interface systems "dynamically" (see p. 00).

THIS IS VERY INTERESTING, AND I THINK NOT FULLY PURSUED IN THIS PAPER. DISPENSING 'WITH THE EXPRESSION ALTOGETHER', IN THE LIMIT, IS NOT JUST DOING THINGS WITHOUT LEVELS, BUT INDEED DOING THEM WITHOUT CATEGORIES, EVEN, HAVING CATEGORIES BE STAGES IN A DERIVATION. THAT WOULD BE RADICALLY DERIVATIONAL. AS FAR AS I CAN SEE, ALSO, THIS SYSTEM IS METHODOLOGICALLY SIMPLER, IN THAT IT MAKES FEWER ASSUMPTIONS (NO LEVELS, NO CATEGORIES). THE OTHER EQUALLY SIMPLE SYSTEM IS FULLY REPRESENTATIONAL (OF THE SORT BRODY WAS AIMING AT). IN BETWEEN THESE TWO EXTREMES, THE SYSTEM WITH LEVELS, CATEGORIES, AND SO ON IS REALLY 'MIXED', HENCE A PRIORI LESS INTERESTING--THOUGH PERHAPS THE RIGHT ONE, AND NECESSARY FOR OTHER REASONS.

A weak> derivational approach assumes that interface levels exist, allowing "post-cyclic" operations that apply to them in whole or in part (deletion of tail of a chain, imposing metrical structure, determining ellipsis and scope, etc.).

THOSE ARE ALL INTERESTING EXAMPLES OF THE WEAKLY DERIVATIONAL SYSTEM, FAMILIAR FROM THE WORK OF EPSTEIN, NUNES, BOSKOVIC, LASNIK, HORNSTEIN, CINQUE, ETC.

There are many options.

AND I THINK THAT'S TO BE KEPT IN MIND: THE RADICALLY DERIVATIONAL SYSTEM ELIMINATES MANY OF THOSE OPTIONS.

With richer set-theoretic assumptions, a recursive definition can be restated as a direct definition, in this case, of the form: E is an expression of L iff ...E..., where ...--... is some condition on E. One might, then, take L to be a direct definition of the set {EXP}, adopting a representational approach>.

THIS IS THE OTHER EXTREME I MENTIONED ABOVE.

Again there are weaker varieties, for example, the assumption that the set of Lfs is given (universally, or by L), with PF derived from LF by some computational procedure. The issue is reminiscent of old questions about morphological processes ("item-and-process" vs. "item-and-arrangement," etc.) and grammatical transformations. Thus, does a transformation map an input to an output structure, or is it an operation on the "output" that expresses properties of the "input"? It is unclear whether these are real questions; on the surface, they look like the question whether 25 = 5<2> or 5 = %25. If the questions are real, they are subtle. They have elicited no little passion over the years, but it is out of place.

{There are many such debates, often with an oddly one-sided character: criticism of a largely unspecified position, with no defense of it on the part of those who are alleged to hold it but who in fact do not see what the issue is. Examples include the "innateness hypothesis," "autonomy of syntax," "formalist" approaches. For an effort to find some significance in the "functionalist-formalist controversy," see Lasnik (1996). For similar attempts with regard to the "autonomy" thesis, see Chomsky (1977). Critics of the "innateness hypothesis" may have in mind issues of modularity and species-specifity [SIC], though that is unclear, since proposals with any substance are highly modular and (so far as is known) species-specific. See references of note <MARCUS>.}

AS YOU CAN IMAGINE, THIS NOTE IS HEAVILY LOADED, AND NOT DIRECTLY RELEVANT TO THE PARTICULAR POINT BEING DISCUSSED NOW (EXCEPT METHODOLOGICALLY, IN WHICH SENSE THE ISSUE IS WELL-TAKEN).

The apparent alternatives seem to be mostly intertranslateable, and it is not easy to tease out empirical differences, if there are any. Surprisingly, there is reason to believe that the questions may be real.

{On conceptual and empirical arguments, with varying conclusions, see, inter alia>, references of note 1 and Chomsky (MP>, 1998).}

The evidence that has been adduced is far from conclusive, and often conflicting. I will adopt the derivational approach as an expository device, though I suspect it may be more than that.

IN FACT, IT IS CENTRAL TO THIS ARTICLE.

If so, that would be a curious and puzzling fact about the nature of the mind/brain.

{Any interpretation of L is computational in some sense, raising difficult and obscure questions

about what this means for a cognitive system. These are not to be confused with problems of processing (parsing, production).}

WE MENTIONED THE PROBLEM OF THE COMPUTER METAPHOR ABOVE, AND SOME OF THE PROBLEMS THAT ARISE WITH REPRESENTATIONS. THE MORE THAT WE GO INTO THE DERIVATIONAL ALTERNATIVES, THE MORE THAT THESE PROBLEMS BECOME SIGNIFICANT, SINCE WE'RE SAYING THAT THE NATURE OF THE SYSTEM IS COMPUTATIONAL IN MORE AND MORE MEANINGFUL WAYS (NOT JUST SYMBOLS, BE THEY REPRESENTED OR NOT, BUT INDEED PHASES OF DERIVATIONS, ALBEIT WITH NO TEMPORAL SIGNIFICANCE AND NO DIRECT CONNECTION TO PARSING OR PRODUCTION).

Suppose that the issue is real, and the derivational approach in fact correct. Then further questions arise. Thus, we might inquire into the complexity of the generative procedure.

THIS IS YET ANOTHER SENSE IN WHICH THE SYSTEM WOULD BE SERIOUSLY COMPUTATIONAL: IT CARES ABOUT COMPUTATIONAL COMPLEXITY, IN A MORE OR LESS STANDARD SENSE.

Such questions have arise over the years, in one or another form. One category concerns "least effort" conditions, which seek to eliminate anything unnecessary: (i) superfluous elements in representations, (ii) superfluous steps in derivations. The tacit assumption is that failure to meet these conditions imposes deviant interpretations, in principle an empirical issue though often not an easy one to resolve. Subcategory (i) involves legibility conditions and convergence ("full interpretation"); (ii) holds that operations are allowed only if there is some reason for them. In the terms we are exploring, reasons are reduced to effects at the interface. Possibilities that have been investigated (if not in these terms) include constraints barring PF-vacuous overt movement and others that seek to limit effect on PF (Procrastinate). An LF counterpart is that covert operations are allowed only if they have an effect on interpretation at LF.

THE IDEA HERE IS THAT THE SYSTEM DOESN'T DEAL WITH ELEMENTS THAT ARE SUPERFLUOUS, HENCE INTRODUCE UNNECESSARY OPERATIONAL COST. THE RATIONALE IS TO ELIMINATE THEM IN ORDER NOT TO CLOG UP THE SYSTEM.

Another category seeks to reduce "search space" for computation: "shortest move/attract," successive cyclic movement (relativized minimality, subjacency), restriction of search to c-command or minimal domains, etc.

THIS TOO MIGHT BE INTERPRETED IN TERMS OF COMPLEXITY: THE SMALLER YOUR SEARCH SPACE IS, THE LEAST OPERATIONAL EFFORT YOU'LL HAVE TO WASTE.

Yet another imposes "local determinability" conditions (barring "look-ahead," "backtracking," or comparison of alternatives).

SIMILARLY HERE, ALTHOUGH THE 'SEARCH SPACE' IN THIS SENSE WOULD BE GOING 'UPWARDS' INTO THE FOLLOWING STEPS IN THE DERIVATION, AND NOT DOWNWARDS INTO WHAT'S ALREADY DERIVATIONALLY COMMITTED, ASSUMING A BOTTOM-UP DERIVATION (ALTHOUGH SEE BELOW FOR AN

ALTERNATIVE).

I will assume the ideas to be generally on the right track, and pursue them further below.

OF COURSE, IT SHOULD BE SAID THAT EVEN IF THE IDEAS ARE GENERALLY ON THE RIGHT TRACK, A) IT IS NOT CLEAR WHY THEY SHOULD BE, WITHIN A COGNITIVE SYSTEM; AND MORE IMPORTANTLY, B) THERE ARE STANDARD MEASURES OF COMPLEXITY IN COMPUTATIONAL SYSTEMS (FORMAL COMPLEXITY THEORY), FOR INSTANCE IN TERMS OF WHETHER A GIVEN PROBLEM IS SOLVABLE IN POLYNOMIAL TIME. NONE OF THE CONCLUSIONS ENTERTAINED HERE REALLY BEAR DIRECTLY ON ANY OF THIS (THE ONLY TRUE EXCEPTION I KNOW OF IS THE DEMONSTRATION BY WEINBERG THAT A SYSTEM WITH MULTIPLE SPELL-OUT, OF THE SORT DISCUSSED BELOW, CAN HANDLE VARIOUS PHRASAL RELATIONS IN POLYNOMIAL TIME, UNLIKE A MORE GLOBAL SYSTEM).

{See Chomsky (1986b) on vacuous movement; Fox (1995, 1998) and Reinhart (1993) on the LF counterpart; Collins (1997) on local determination. And much other work.}<FXRN = footnote>

Some of these notions have analogues in formal complexity theory.

AS I SAID, VERY FEW OF THOSE. IN CONTRAST:

Most are the kinds of intuitive ideas about "operative complexity" that enter commonly into the cognitive sciences

{For example, integrated action/perception models motivated by computational savings over construction of the detailed properties of a presented scene. For review, see Clark (1998).}

and design considerations generally.

THAT'S FINE, BUT DESIGN IS AN OPEN QUESTION--THERE'S NO GENERAL THEORY OF DESIGN (WHICH IS OFTEN NOT REALIZED BY NEO-DARWINISTS; BUT SEE DAVID BERLINSKI'S WORK ON *ACTUAL* DESIGN OF ARTIFACTS).

Suppose automobiles lacked fuel storage, so that each one had to carry along a petroleum processing plant. That would add only bounded "complexity," but would be considered rather poor design.

IT DEPENDS ON HOW LARGE THE PETROLEUM PLANT IS, HOW MUCH PETROLEUM IT PRODUCES, AND SO FORTH. A HOT TOPIC IN SPACECRAFT DESIGN IS PRECISELY TO FIND OUT WAYS OF PROPULSION THAT DO NOT NEED REFUELING (THE ULTIMATE 'DILITHIUM CHAMBER'...). THE POINT IS, WITHOUT A FORMAL THEORY OF COMPLEXITY, OR SOME INDEPENDENT METRIC, OR GOOD SOLID SUPPORT FROM THE HARD SCIENCES, OPTIMALITY IN DESIGN MAY WELL JUST BE REIFYING OUR FAVORITE ACCOUNT.

Something similar might well be true for language. Let's consider a few such proposals, beginning with conventional ones and proceeding to others that are more controversial; it's worth bearing in mind, however, that the logic is similar throughout.

UG makes available a set {F} of features (linguistic properties)

THIS IS JUST A PARENTHETICAL, BUT AN INTERESTING CONSEQUENCE IS THIS: IF FEATURES ARE PROPERTIES, WHAT ENTITIES ARE THEY PROPERTIES OF? FURTHERMORE, ARE WE PREPARED TO SAY THAT Chl MANIPULATES PROPERTIES, INDEPENDENT OF THE ENTITIES THEY ARE PROPERTIES OF? IF NOT, WHAT IS THE MINIMAL UNIT THAT Chl OPERATES ON?

and operations C<HL> (the computational procedure for human language) that access {F} to generate expressions. The language L maps {F} to a particular set of expressions {EXP}. Operative complexity is reduced if L makes a one-time selection of a subset [F] of {F}, dispensing with further access to {F}. It is reduced further if L includes a one-time operation that assembles elements of [F] into a lexicon LEX, with no new assembly as computation proceeds.

OPERATIONAL COMPLEXITY IS A CENTRAL THEME OF THIS PAPER. I SUPPOSE THE SUGGESTION HERE IS THAT THE REASON A LANGUAGE USES A GIVEN SUBSET OF THE FEATURES AVAILABLE, AND FURTHERMORE ARRANGES THOSE INTO A LEXICON, HAS TO DO WITH COMPLEXITY. CERTAINLY, THE POINT IS GRANTABLE FOR WORDS IN A LEXICON (FIXED COMBINATIONS OF FEATURES, PRESUMABLY), BUT IS HARDER TO GRANT AS A RATIONALE FOR GIVEN LANGUAGES CHOOSING ONLY GIVEN FEATURES (ASSUMING THAT'S EVEN FACTUALLY CORRECT, A HARD THING TO ESTABLISH). IN THE LIMIT, NOTE, THE ASSUMPTION IS MILDLY WORFIAN (NOT STRONGLY, SINCE PRESUMABLY THE VOCABULARY IS UNIVERSAL).

On these (conventional)

I DON'T KNOW WHETHER THE WORD 'CONVENTIONAL' IS USED TO MEAN 'OBVIOUS' OR RATHER 'TRADITIONAL' OR EVEN 'WITH NO MOTIVATION, AS A MERE CONVENTION THAT COULD HAVE BEEN DROPPED'.

assumptions, acquiring a language involves at least selection of the features [F], construction of lexical items LEX, and refinement of C<HL> in one of the possible ways

NOTE THAT PARAMETER SETTINGS ARE SEEN AS DIFFERENT FROM CHOICE OF FEATURES FROM THE UNIVERSAL VOCABULARY.

-- parameter setting.

{The properties of features and assembly form a large part of the subject matter of traditional and modern linguistics; I will put these topics aside here, including questions about organization of assembled features within a lexical item LI. Also left to the side is the question whether LI is assembled in a single operation or at several stages of the derivation, as in Distributed Morphology (Halle and Marantz 1993). Rephrasing of the account just given in these terms is straightforward. Recall that L is a state> of FL; state changes, of course, modify the lexicon.}<DM = footnote>

I DON'T FULLY UNDERSTAND THE LAST SENTENCE IN THE FOOTNOTE. THE REST OF THE FOOTNOTE IS STRAIGHTFORWARD, AND DESERVES TO BE KEPT IN MIND

WHEN PURSUING THE PROGRAM--THE DECISION HERE IS ARBITRARY.

One could offer a conceptual argument that conventional assumptions are mistaken, on the grounds that a theory lacking certain concepts (here [F], LEX, and the operations forming [F] and LEX) is better than an otherwise identical one that employs them.

OF COURSE THAT WOULD BE AN ARGUMENT BASED ON METHODOLOGICAL SIMPLICITY.

But if operative complexity matters, the argument loses force. Conceptual arguments can be given either way. The issues are empirical, and can be settled only by investigating consequences of alternative conceptions, considered so obvious in this case that the question has not arisen.

I'M NOT REALLY SURE THE REASON QUESTIONS HAVE NOT ARISEN IS BECAUSE THE ISSUES ARE OBVIOUS. I DON'T KNOW OF ANY ARGUMENT TO THE EFFECT THAT WORDS OR SUB-SETS OF UNIVERSAL VOCABULARIES EXIST BECAUSE OF THAT, ALTHOUGH THE TAKE IS NOT IMPLAUSIBLE.

We assume, then, that a language L maps ([F], LEX) to {EXP}. The next natural simplification would be to reduce access to the domain ([F], LEX) of L. Consider [F]. In the computation of LF -- what we may call narrow syntax> -- it seems that [F] is not accessed, only LEX (and features of its items).

THAT IS, WE DON'T ACCESS FEATURES DIRECTLY.

The restriction does not extend to phonology, however: features are introduced in the course of computation, and in different ways for different languages, whatever approach one takes to computation of PF.

AGAIN, NON-TRIVIAL ISSUES ARISE HERE ABOUT WHAT IT IS THAT THE SYSTEM MANIPULATES, BUT THE POINT IS GENERALLY WELL-TAKEN (PHONOLOGY DOES SEEM TO MODIFY SUB-LEXICAL STUFF).

Keeping to narrow syntax then, we may take C<HL> to be a mapping of LEX to the Lfs of {EXP}. Is it also possible to reduce access to LEX, the second component of the domain of L?

OF COURSE, THE ANSWER TO THIS RHETORICAL QUESTION IS 'YES', BUT BE AWARE OF THE FACT THAT 'ULTIMATE' REDUCTION WOULD CREATE A WORD PER MESSAGE, A SORT OF NAME FOR DIFFERENT EVENTS. IN OTHER WORDS, WHY DON'T LANGUAGES WORK IN SUCH A WAY THAT 'JOHN DESTROYED ROME' IS 'CALLED' 'BUM', AND 'MARY ATE PEANUTS' IS CALLED 'BAM', AND SO ON. THAT WOULD BE A RATHER DRASTIC REDUCTION IN TERMS OF ACCESS TO THE LEXICON, BUT SOMEHOW, INTUITIVELY, WE DON'T WANT TO GO THAT FAR. (NOTICE THAT SAYING IT IS BECAUSE THE LEXICON IN QUESTION WOULD ONLY ALLOW FOR A FINITE NUMBER OF MESSAGES DOESN'T EXPLAIN ANYTHING; YOU WOULD HAVE TO SHOW WHY SUCH A SYSTEM DIDN'T ARISE, IF IT IS IN SOME SENSE COMPUTATIONALLY VALID OR EVEN OPTIMAL; IT APPEARS THAT YOU ALSO WANT TO INTRODUCE METRICS OF 'EXPRESSIVENESS' OR EVEN 'PLASTICITY', BUT GOOD LUCK ON ALL OF THAT--IT'S GETTING TO BE

DANGEROUSLY CLOSE TO THE SORTS OF ISSUES THEY TALK ABOUT IN 'COMMUNICATION THEORY'.) IN THE ABSENCE, HOWEVER, OF A THEORY OF THE LEXICON OR A THEORY EVEN OF HOW LARGE A LEXICON COULD BE (FOR REASONS OF SIZE, EXPRESSIVENESS, PLASTICITY, AND THE LIKE), THE POINT IS VERY VAGUE.

The obvious proposal is that derivations make a one-time selection of a lexical array> LA

THIS IS CERTAINLY A PROPOSAL, BUT I DON'T SEE WHY IT IS SO OBVIOUS. AFTER ALL, THE LEXICON IS 'IN THE HEAD', CLEARLY. ONCE THE INFORMATION IS PRESENT, TO SAY THAT MAKING IT 'ACTIVE' AS A WHOLE IS COMPUTATIONALLY MORE COMPLEX THAT HAVING IT THERE IN MEMORY IS, TO SAY THE LEAST, A SOPHISTICATED PROPOSAL TO MAKE--NOTHING OBVIOUS. OF COURSE, IF THE LATTER ASSUMPTION HOLDS, THEN THE SIZE OF THAT 'ACTIVE' PART MATTERS, AND HAVING A LEXICAL ARRAY IS GOOD.

from LEX,

{Or, if we distinguish independent selections of a single lexical item, a numeration NUM (as in MP>),

NOTE THIS IS ALL THAT IS MEANT BY A 'NUMERATION', A SET OF LEXICAL TOKENS, NOT LEXICAL TYPES.

an extension I will put aside until it becomes relevant.}<NUMER = footnote>

then map LA to expressions, dispensing with further access to LEX. That simplifies computation far more than the preceding steps.

AGAIN, ONLY GIVEN CERTAIN ASSUMPTIONS ABOUT THE ARCHITECTURE (HAVING THE LEXICON 'ACTIVE' AS OPPOSED TO 'JUST COMMITTED TO MEMORY').

If the derivation accesses the lexicon at every point, it must carry along this huge beast,

AGAIN, THIS IS MISLEADING. THE 'HUGE BEAST' IS THERE, AFTER ALL, IN THE HEAD. THE QUESTION IS WHETHER THE COMPUTATIONAL SYSTEM ADDS SOME COST BECAUSE OF HAVING TO CONSULT WITH THE BEAST...

rather like cars that have to replenish fuel supply constantly.

{It would not suffice to say that constant memory can be accessed throughout the derivation. The lexicon is a distinct component of memory; for C<HL>, our beliefs about the stars don't matter, but the lexical properties of "star" do. However hard it may be to make the distinction properly, there is good reason to believe that it is real.}

SO OF COURSE THIS IS WHERE THE CRUCIAL ASSUMPTION IS, ABOUT ARCHITECTURE. WHICH IS WHY I'M SAYING THERE'S NOTHING OBVIOUS ABOUT THE PROPOSAL, EVEN IF IT MAY BE REASONABLE, GIVEN WHAT'S KNOWN ABOUT MEMORY AND COMPUTATIONAL PROCEDURES MORE GENERALLY.

Derivations that map LA to expressions require lexical access only once, and thus reduce operative complexity in a way that might well matter for optimal design.

FOR WHAT IT'S WORTH, NOTE THAT A SYSTEM WITH D-STRUCTURE DIDN'T FACE THIS ISSUE, OBVIOUSLY (ONCE D-STRUCTURE WAS ASSEMBLED, IT CARRIED YOU TO A VERY LIMITED SET OF POSSIBLE LFS). HERE THE ARGUMENT BEING MADE MAKES THE LEXICAL ARRAY DIFFERENT FROM D-STRUCTURE IN TWO RESPECTS. FIRST, D-STRUCTURE WAS, AS ITS NAME INDICATED, STRUCTURED; THE ARRAY PURPORTS TO BE UNSTRUCTURED (WE'LL SEE WHETHER THAT'S TRUE--IF IT IS A NUMERATION, THE ARRAY IS ALREADY SOME SORT OF CONSTRUCTION (A SET OF TOKENS) FROM THE LEXICON, HENCE NOT ENTIRELY UNSTRUCTURED). SECOND, D-STRUCTURED WAS A SUBSTANTIVE SYSTEM, WITH SYMBOLS AND INTERACTIONS AMONG THEM OF A GIVEN SORT, AND A GIVEN GLOBALITY; THE LEXICAL ARRAY DOES HAVE SOME GLOBALITY (AFTER ALL, THE ENTIRE SET OF ITEMS THAT WILL HAVE TO BE USED IN A DERIVATION) BUT DOESN'T REALLY HAVE MUCH OF A SUBSTANTIVE CHARACTER. OF COURSE, WHETHER THAT'S COMPLETELY THE CASE WILL DEPEND ON WHAT GOES IN THE ARRAY (TYPES OR TOKENS? CLASSES OF TOKENS? ETC.) AND WHAT KINDS OF INTERACTIONS IT ALLOWS (IDEALLY NONE, IF IT IS JUST A COMPUTATIONAL DEVICE TO REDUCE COMPLEXITY--BUT LET'S SEE WHETHER THIS IS ENTIRELY TRUE).

Again, conceptual arguments can be given either way, but they carry little weight. The questions are empirical. Investigating them, we can hope to discover whether (and if so how) what might reasonably be considered complexity/economy considerations enter into language design. If FL operates with the economy principles just reviewed, then a language L follows the procedure (I,II) of (3) to specify the language (apart from parameter setting), then applies (III,IV) to derive a particular EXP:

(3)

(I) Select [F] from the universal feature set {F}

(II) Select LEX, assembling features from [F]

(III) Select LA from LEX

(IV) Map LA to EXP, with no recourse to [F] for narrow syntax

IT SHOULD BE KEPT IN MIND THAT ONLY (III) AND (IV) ARE STRICT DERIVATIONAL PROCEDURES, IN THE TRADITIONAL SENSE OF THE WORD 'DERIVATION'. OF COURSE, THERE'S NO REASON TO EXPECT THAT THE MECHANISM OF FL DOESN'T INVOKE OTHER PROCEDURES THAT DO NOT FALL WITHIN A DERIVATION, AND ARE THERE JUST TO REDUCE THE COMPLEXITY OF DERIVATIONS. IN EFFECT, NOTE, WHAT (I) AND (II) DO IS ARTICULATE THE LEXICON IN A CERTAIN WAY. EVEN WITHIN THE NARROW CONFINES OF A COMPUTATIONAL/DERIVATIONAL THEORY (WHICH MUST BY DEFINITION INCLUDE A LEXICON AND A COMPUTATIONAL PROCEDURE) THERE IS NO

REASON ONE SHOULDN'T ARTICULATE THE LEXICON IN WHICHEVER WAY ONE PLEASES. OF COURSE, THE QUESTION THEN IS WHY IT IS ARTICULATED *THAT* WAY, AND HERE CHOMSKY'S CONJECTURE IS IN TERMS OF COMPUTATIONAL COMPLEXITY. (AS I SAID, OTHER DIFFICULT ISSUES PROBABLY ALSO ARISE, SUCH AS LEXICON SIZE, EXPRESSIVENESS, PLASTICITY, GOD KNOWS); AGAIN, EVEN IF THOSE CONSIDERATIONS ARE COMPLETELY VAGUE AND SOME OF THEM PERHAPS ENTIRELY OFF THE MARK, THE POINT IS THAT THE THEORY OF FL SHOULD BE ABLE TO POSE SUCH QUESTIONS, ANSWERING THEM ONE WAY OR THE OTHER, FOR INSTANCE IN TERMS OF ARTICULATING THE LEXICON. AND INCIDENTALLY, SIMILAR ISSUES ARISE FOR THE COMPUTATIONAL SIDE (OR NARROW SYNTAX) WHICH HAVE NOTHING TO DO WITH LEGIBILITY CONDITIONS, AND FALL WITHIN WHAT WE USUALLY CALL 'VIRTUAL CONCEPTUAL NECESSITY', WHICH IS ANOTHER WAY OF TALKING ABOUT FORMAL RESTRICTIONS ON THE COMPUTATIONAL SYSTEM; OF COURSE, THE NARROW MINIMALIST THESIS DOESN'T ALLOW YOU TO TALK ABOUT THOSE UNLESS THEY REALLY FALL ON THEIR OWN WEIGHT, SINCE THE COMPUTATIONAL SYSTEM IS NOT THOUGHT TO HAVE A LIFE OF ITS OWN--IT'S THERE JUST TO MEET LEGIBILITY REQUIREMENTS. STILL, THAT LEAVES SOME ROOM FOR MANEUVER, AS WE WILL SEE IN TERMS OF COMPLEXITY, BUT PERHAPS IN OTHER LESS UNDERSTOOD TERMS AS WELL--SEE BELOW THE 'SLAVING PRINCIPLE' ON COMPLEX SYSTEMS.)

We return to further steps along the same path, but let us first look more closely at general properties of LEX and narrow syntax (the recursive part of L). First, what operations enter into this component of C<HL>? One is indispensable in some form for any language-like system: the operation Merge>, which takes two syntactic objects (",ß) and forms K(",ß) from them.

WELL, HOLD ON. REMEMBER, WE'RE TRYING TO DERIVE PROPERTIES OF THE SYSTEM FROM EITHER LEGIBILITY CONDITIONS OR VIRTUAL CONCEPTUAL NECESSITY. THERE IS MUCH PACKED INTO THE NOTION 'ANY LANGUAGE-LIKE SYSTEM' HERE. IF WE MEAN LITERALLY HUMAN LANGUAGE, THEN WE'RE BEING CIRCULAR. IF WE MEAN OTHER COMPLEX SYSTEMS USING SOMETHING LIKE SYMBOL TOKENS, THEN I DON'T SEE IT AS INDISPENSABLE THAT (I) THE COMBINATIONS SHOULD BE BINARY, (II) THEY SHOULD ARRANGE THEMSELVES IN 'TREE-LIKE' FASHION, AND (III) THERE SHOULD BE ANY NON-TRIVIAL COMBINATIONS AT ALL. (I) IS OBVIOUS, AND EVEN CHOMSKY, I THINK, WOULD GRANT THAT YOU WANT THAT TO FOLLOW FROM SOMETHING ELSE (E.G. KAYNE'S LCA--IT DOES FOLLOW FROM THAT, ALTHOUGH WHETHER THE LCA IS ULTIMATELY RIGHT IS ALSO EMPIRICAL). (II) SHOULD ALSO BE OBVIOUS, THE MINUTE YOU THINK OF MOLECULES OR PROTEIN FOLDING. ARE THOSE LANGUAGE-LIKE? WELL... YOU CAN EXPRESS THEM COMPUTATIONALLY, THAT'S FOR SURE. SIMILARLY, (III) POSES NON TRIVIAL QUESTIONS, PARTICULARLY BECAUSE SYSTEMS EXIST OUT THERE WHICH, IN ESSENCE, ARE MERE CONCATENATIONS, THUS NOTHING REMOTELY IN NEED OF A PROCEDURE LIKE 'MERGE' (STARTING WITH DNA STRANDS). MY POINT IS, WE SHOULDN'T BE TOO FAST AND LOOSE ABOUT 'MERGE'. SURELY WE NEED IT IN LINGUISTICS, BUT WE SHOULD HAVE TO DEMONSTRATE THAT IT FOLLOWS EITHER FROM INTERFACE

PROPERTIES OR NATURAL INTERACTIONS IN THE SYSTEM. YOU MAY RECALL THAT THERE IS AN ATTEMPT TO HAVE 'MERGE' FOLLOW THAT WAY IN CHAPTER IV. I'M NOT SURE IT'S CONVINCING, BUT NONETHELESS THE POINT IS SOME HOMEWORK NEEDS TO BE DONE HERE, TO SHOW THAT 'MERGE' IS EITHER A NECESSARY OR AT ANY RATE OPTIMAL SOLUTION THAT THE SYSTEM HAS FOUND TO THIS, THAT, OR THE OTHER.

A second is an operation we can call Agree>, which establishes a relation (agreement, Case-checking) between an LI " and a feature F in some restricted search space (its domain>).

FASTEN YOUR SEATBELTS, FOR THIS IS NEW. IT'S ONLY SOMEWHAT NEW, BUT NONETHELESS, IT WASN'T THERE IN CHAPTER 4 (ALTHOUGH FEATURE ATTRACTION CAME CLOSE, AS WE WILL SEE). I SHOULD SAY, BY THE WAY, THAT IT IS NOT TOTALLY CLEAR TO ME WHAT KIND OF OBJECT 'AGREE' PRODUCES. THAT IS, 'MERGE' PRODUCES A FAMILIAR PHRASE-MARKER, AND 'MOVE' (HOWEVER YOU CARE TO UNDERSTAND IT) PRODUCED LESS FAMILIAR, YET ULTIMATELY CLEAR SETS OF PHRASE-MARKERS OR CHAINS. HOWEVER, WE WILL HAVE TO PONDER VERY CAREFULLY WHAT EXACTLY 'AGREE' GENERATES. WE WILL RETURN TO THAT.

Unlike Merge, this operation is language-specific, never built into special-purpose symbolic systems and apparently without significant analogue elsewhere.

OF COURSE, THAT'S THE GUIDING INTUITION BEHIND 'MOVE', ALSO. DIFFERENTLY PUT, WHETHER WE DEAL WITH MOVE, ATTRACT, OR AGREE, THERE IS A PROPERTY IN THE SYSTEM, CALL IT XTRA, THAT APPEARS TO BE EXTREMELY PECULIAR. THIS ONE, CHOMSKY HAS NEVER EVEN CLAIMED TO HAVE GOOD REASONS FOR. QUITE THE OPPOSITE, HE HAS OFTEN COMMENTED ON XTRA AS A QUASI-GLITCH IN THE SYSTEM, HOPEFULLY RELATED TO OTHER SUCH GLITCHES: MORPHOLOGY, UNINTERPRETABLE FEATURES, LINGUISTIC VARIATION, AND PERHAPS OTHERS.

We are therefore led to speculate that it relates to the design conditions for human language.

A FAIRER WAY OF SAYING THAT WOULD BE THAT WE ARE LED TO SPECULATE THAT XTRA RELATES EITHER (I) TO AN AS OF YET NOT UNERSTOOD LEGIBILITY CONDITION OR (II) TO DESIGN CONDITIONS. (I) IS EQUALLY PLAUSIBLE, AND HAS BEEN PURSUED FOR INSTANCE BY HORNSTEIN. OF COURSE, IT IS HARD TO SEE HOW THAT GENERAL LINE (A) GIVES YOU A PRECISE REASON FOR THE EXISTENCE OF XTRA, AS OPPOSED TO ANY OTHER PROPERTY, OR (B) THE CONCLUSIONS EXTEND FROM, ROUGHLY, A'-MOVEMENT AND THE "SEMANTICALLY RELATED" TRANSFORMATIONS TO ALL OTHERS--ALTHOUGH HORNSTEIN HAS SOME SPECULATIONS ABOUT THAT. BUT IN ANY CASE, THESE ARE MATTERS OF FACT, NOT OF PRINCIPLE, AND A CLEVER ENOUGH ANALYSIS MIGHT WELL GET YOU A DERIVATION OF XTRA FROM (I). BUT FOR WHATEVER REASON, CHOMSKY PREFERS TO EXPLORE (II). I TEND TO AGREE WITH THAT GUT FEELING, BUT IT IS ESSENTIALLY THAT AND NOTHING MORE. SINCE THE DEVIL IS IN THE DETAILS, WE'LL HAVE TO SEE WHETHER CHOMSKY'S ACCOUNT OF

XTRA FROM (II) IS CONVINCING, AND OTHERWISE WHETHER THERE ARE OTHER PLAUSIBLE DESIGN CONDITIONS TO DERIVE XTRA FROM.

A third operation is Move>, combining Merge and Agree.

HERE WE'RE GETTING TECHNICAL, IN IMPORTANT WAYS. WHAT I WAS VAGUELY CALLING XTRA BEFORE IS NOW PRECISELY SEPARATED. MOVE IS MERGE+AGREE, WHICH RECALLS THE CHAPTER 4 SYSTEM WHERE MOVE WAS MERGE+ATTRACT (PLUS SOMETHING ELSE HAVING TO DO WITH MORPHOLOGICAL REALIZATION, WHICH AS FAR AS I KNOW WILL BE NECESSARY HERE TOO).

The operation Move establishes agreement between " and F and merges P(F) to "P, where P(F) is a phrase determined by F (perhaps but not necessarily its maximal projection) and "P is a projection headed by ".

LET'S SEE THIS IN A TREE:

MOVE:

(II) MERGE:

(I) AGREE:

"P

^

"P [...F...]<--|--> "P

/ \ P(F) / \

" ... " ...

. / \

. P(F)

. [...F...]

.....^

P(F) becomes SPEC-". Let us refer to Move of P to SPEC-N as A-movement, where N is an agreement feature (N-feature); other cases are A'-movement.

OF COURSE THE ENTIRE BITE, THEN, IS IN UNDERSTANDING WHAT THOSE FEATURES ARE, WHY THE SYSTEM HAS THEM, WHY THEY APPEAR ONLY IN THE CATEGORIES THEY APPEAR, AND SO ON. RECALL THAT IN CHAPTER 3 THE DISTINCTION BETWEEN A AND A' MOVEMENT WAS TAKEN TO DEPEND ON THE NOTION 'L-RELATEDNESS', ALSO RATHER UNCLEAR. WE WILL RETURN TO ALL

THIS AS WE GO ALONG.

Plainly Move is more complex than its subcomponents Merge and Agree, or even the combination of the two, since it involves the extra step of determining P(F) (generalized "pied piping"). Good design conditions would lead us to expect that simpler operations are preferred to more complex ones, so that Merge or Agree (or their combination) preempt Move, which is a "last resort," chosen when nothing else is possible. Preference of Agree over Move yields much of the empirical basis for Procrastinate and has other consequences, as do the other preferences.

THAT IS, IN THE CHAPTER 3 SYSTEM YOU HAD TO SAY THAT THE SYSTEM PROCRASTINATES (TRIES TO DELAY MOVEMENT) IF IT CAN. ALREADY IN THE CHAPTER 4 SYSTEM IT BECAME POSSIBLE TO SAY THAT MOVE EMPLOYED MORE OPERATIONS THAN OTHER COMPUTATIONAL PROCESSES, HENCE WAS ALWAYS TRUMPED BY EITHER MERGE OR SIMPLE ATTRACTION OF FEATURES (WITH NO ANCILLARY MERGE). THAT ARGUMENT CARRIES THROUGH TO AGREE. WE WILL SEE OTHER CONSEQUENCES OF AGREE BELOW.

{Much current work allows Move to raise " to SPEC-V even when Merge or Agree would be options; see Lasnik (1995c) and sources cited. An interesting question is whether the evidence for this conclusion can be incorporated within the more restricted framework envisaged here under a proper interpretation of the qualification "when possible."}

SO THERE YOU HAVE A THESIS TOPIC... WE WILL RETURN TO THESE SORTS OF EXAMPLES IN THE SEMINAR.

Let us turn next to the lexicon LEX, adopting some fairly common assumptions along with more controversial ones, and keeping to simple cases. LIs fall into two main categories: substantive and functional; we are concerned now mainly with the latter. Take the core functional categories> CFCs to be C (expressing force/mood), T (tense/event structure), and v>, the "light verb" head of transitive constructions.

IT MIGHT BE SLIGHTLY MISLEADING TO BLAME 'EVENT' STRUCTURE ON T. THERE ARE TWO SIDES TO EVENT STRUCTURE. ONE HAS TO DO WITH SO CALLED 'EXTERNAL' ASPECT (WHETHER AN ACTION IS IN PROGRESS OR COMPLETED, AND SO ON) AND ALSO THE ANCHORING OF TEMPORAL AND OTHER ADVERBS IN THE TIME OF THE EVENT, AND SIMILAR ISSUES; THAT DOES SEEM TO BE RELATED TO T, IN SOME SENSE. BUT EVENT STRUCTURE ALSO ARISES IN WHAT YOU MAY THINK OF AS THE INTERNAL MAKE-UP OF EVENTS, THEIR 'INTERNAL' ASPECT (WHETHER THE VERB IS A STATE OR AN EVENT, ETC.), THE PROPER LEXICAL ENTAILMENTS THEY SANCTION, AND SO ON. ALL THAT SEEMS TO BE DIRECTLY RELATED NOT TO T, BUT RATHER TO v, AS WE WILL SEE IN SOME DETAIL LATER ON. ALSO, WHETHER MOOD IS A PROPERTY OF C, OR OF T, OR OF SOME INTERACTION BETWEEN BOTH IS SOMETHING THAT WILL HAVE TO BE ESTABLISHED. NEEDLESS TO SAY, THOUGH, THIS IS ALL EMPIRICAL.

All CFCs may have N-features (obligatory for T, v>).

MAKE NO MISTAKES: THIS IS A HUGE ASSUMPTION. IN FACT, TWO

ASSUMPTIONS: (I) THAT CFC HAVE THOSE FEATURES, AS OPPOSE TO OTHERS; (II) THAT THEY ARE OBLIGATORY IN THE CASE OF T AND v. A MINIMALIST THEORY OF ALL THIS HAS TO HAVE THOSE ASSUMPTIONS DERIVED.

These are uninterpretable, constituting the core of the systems of (structural) Case-agreement and "dislocation" (Move).

OBVIOUSLY, A THIRD ASSUMPTION--THEIR UNINTERPRETABILITY. MIND YOU, THIS ALL SEEMS RIGHT, BUT IT IS NOT ENOUGH TO GET THE FACTS, GIVEN THE STANDARDS WE HAVE SET FOR OURSELVES.

Neither T nor v> assigns inherent Case; other light verbs may, as may substantive categories.

ALSO A HUGE TOPIC, STARTING WITH WHAT ON EARTH INHERENT CASE IS. THE COMMENT ABOUT OTHER LIGHT VERBS BEING ABLE TO ASSIGN INHERENT CASE IS THERE JUST IN CASE DATIVE MAY BE INHERENT, AT LEAST IN SOME LANGUAGES. BUT THIS ISSUE IS EXTREMELY MURKY AND WE'LL RETURN TO IT.

{I am putting aside many questions concerning the substantive/functional distinction, adopting it only for heuristic purposes.

IT MAY BE WORTH COMMENTING ON THIS, FOR THE ISSUE IS EXTRAORDINARILY IMPORTANT AND POORLY UNDERSTOOD. BEING DELIBERATELY NAIVE ABOUT IT, WE KNOW, FOR INSTANCE, THAT THE SUBSTANTIVE AND FUNCTIONAL DISTINCTION IS AS ROBUST AS IT GETS IN TERMS OF BRAIN REPRESENTATION, APHASIC BEHAVIOR, LANGUAGE IMPAIRMENT, AND SO ON. FOR INSTANCE, GENIE HAD, AS FAR AS ANYBODY COULD TELL, NO FUNCTIONAL SYSTEM, ALTHOUGH SHE COULD ACQUIRE A SUBSTANTIVE ONE. INDEED BICKERTON HAS SPECULATED WITH THE POSSIBILITY THAT PRE-HOMINIDS DIDN'T HAVE A FUNCTIONAL SYSTEM, OR EVEN THAT OTHER MAMMALS MIGHT BE SOMEWHAT CAPABLE OF AT LEAST COMPREHENDING IMPORTANT ASPECTS OF THE SUBSTANTIVE SYSTEM, WITHOUT EVER GETTING ANYWHERE NEAR THE FUNCTIONAL ONE. ON THE FACE OF IT, THIS SHOWS THAT THE FUNCTIONAL SYSTEM CAN HAVE ABSOLUTELY NOTHING TO DO WITH LEARNING, SINCE IF ANYTHING WHAT IS TO BE LEARNED WITH REGARDS TO IT (THAT A COUPLE OF NOTIONS HAVE OR DO NOT HAVE A SMALL WORD AS THEIR ASSOCIATED PRONUNCIATION) SEEMS TRIVIAL IN COMPARISON WITH WHAT HAS TO BE LEARNED FOR THE SUBSTANTIVE SYSTEM (OFTEN ARCANE PRONUNCIATIONS, INDEFINITELY MANY OF THEM, FOR WHAT IN SOME RESPECT LOOK LIKE ARBITRARILY COMPLEX NOTIONS--I'M EXAGGERATING TO MAKE THE POINT). SO IT SEEMS AS IF THE FUNCTIONAL SYSTEM IS SOME KIND OF ARRAY OF FUNDAMENTAL DIMENSIONS THAT FL IS BASED ON, SO THAT IF YOU REMOVE IT FROM TRIGGERING EXPERIENCE (AS IN GENIE'S CASE) THE ENTIRE SYSTEM COLLAPSES. THEN THE TASK AHEAD WOULD BE TO DISENTANGLE THE PROPERTIES OF THIS SYSTEM. WHY IS IT ARTICULATED AROUND C, T, v, OR WHATEVER ITS COMPONENTS ARE? WE ESSENTIALLY HAVE NO CLUE ABOUT ANY OF THIS. NOTE, CLAIMING THAT THOSE ARE, SAY, THE NOTIONS YOU NEED FOR 'FREGEAN' INTERPRETATION (OR ANY SUCH CLAIM) MISSES THE POINT IN

TWO WAYS. (I) IT'S NOT LIKE GENIE DIDN'T HAVE FREGEAN INTERPRETATION--SHE APPARENTLY KNOWS TRUE FROM FALSE AND EVEN RIGHT FROM WRONG, HAS COMPLEX INTENTIONAL POSITIONS ON EVENTS AROUND HER, AND SO FORTH; SHE JUST DOESN'T HAVE OUR FL. (II) FREGEAN AND SIMILAR SYSTEMS OF INTERPRETATION ARE BASED ON AN ANALYSIS OF THE LOGICAL PROPERTIES OF LANGUAGE, THEY DO NOT CONSTITUTE AN EMPIRICAL HYPOTHESIS ABOUT HOW THE HUMAN MIND IS ARTICULATED; IT COULD BE OF, COURSE, THAT THIS IS INDEED HOW LOGICAL FORMS WORK IN MENTALESE OR WHATEVER, BUT THAT HAS TO BE ESTABLISHED, AND MEANWHILE WE DO NOT HAVE AN UNDERSTANDING AS TO WHY THE CFCS ARE THE ONES WE THINK WE HAVE. OF COURSE, SIMILAR QUESTIONS ARISE ABOUT THE SUBSTANTIVE SYSTEM, STARTING WITH WHY WE CARVE UP THE WORLD IN TERMS OF VERBS AND NOUNS, AND SO ON. THE LOGICAL TRIVIALITIES CONCERNING THIS ARE OFTEN HARD TO BELIEVE. FOR INSTANCE, WHAT WOULD GO WRONG WITH A SYSTEM THAT USED NOUNS TO DESIGNATE COMPLEX EVENTS AND VERBS TO DESIGNATE ITS PARTICIPANTS? FOR EXAMPLE, THE WAY TO SAY 'THE MAN DIED' WOULD BE SOMETHING LIKE THIS: 'MANED THE DYING' (IN MORE 'HUMAN' TERMS, 'THE DYING WAS MANNED'). NEEDLESS TO SAY, THERE WOULD BE CERTAIN THINGS YOU COULDN'T SAY GIVEN THAT SYSTEM AND THE SYNTAX WE HAVE; FOR INSTANCE, YOU COULDN'T THEN EXPRESS TRANSITIVE ACTIONS. BUT YOU WOULD BE ABLE TO EXPRESS NEW THINGS, FOR INSTANCE GENERALIZED QUANTIFICATION OVER NOMINAL EXPRESSIONS, WHICH YOU NOW CANNOT HAVE--'MOST PEOPLE'S DYINGS IMPRESS ME' DOESN'T MEAN 'THAT MOST PEOPLE DIE IMPRESSES ME'). THE POINT IS, WE DON'T HAVE A CLUE ABOUT ANY OF THIS, AND WE DON'T EVEN OFTEN TALK ABOUT IT, KEEPING A STRAIGHT SCIENTIFIC FACE WHEN WE SAY 'WELL, THE SYSTEM MANIPULATES OH, NOUNS AND VERBS.'

Also omitted is the D head of DP, which seems to belong to a different system, more complex verbal constructions, and the question whether nontransitive ones have a light verb head.

DITTO.

Some might, e.g., seem>, which c-commands the experiencer SPEC even in non-raising languages like English, either because it selects a light verb with this SPEC or raises from that position to a light verb.

THIS REFERS TO THE PUZZLING PROPERTIES OF EXPERIENCERS ACROSS LANGUAGES, AND WHETHER THEY INVOLVE SOME VERB MOVEMENT--WE WILL RETURN TO DISCUSS THESE MATTERS.

Ignored as well are the "peripheral" systems outside TP; I will use C and T as surrogates for richer systems. On these matters see Rizzi (1995), Cinque (forthcoming), and many other studies on the CFC systems and others. The concepts inherent/structural Case are understood in the sense of Chomsky (1981, 1986a): theta-related vs. structurally determined.}<LIGHT = footnote> Consider the selectional properties of CFCs, beginning with semantic (s-)selection.

{In the sense of Pesetsky (1982), modifying ideas of Grimshaw (1979).}

Assume that C can be unselected (root), while v> and T cannot.

I DON'T WANT TO KEEP REPEATING MYSELF, BUT I HOPE YOU SEE HOW THESE FURTHER ASSUMPTIONS FALL INTO THE SAME CATEGORY OF ISSUES I HAVE BEEN POSING, AS DO THE FOLLOWING SENTENCES:

C is selected by substantive categories, v> only by a functional category. T is selected by C or V. If selected by C it has a full complement of N-features; if by V it is defective> (T<def>). C selects T, while T and v> select verbal elements. v> may also select a nominal phrase NP/DP as its external argument> EA = SPEC-v>.

{Strengthening "may" to "must" stipulates part of Burzio's generalization; the rest should follow from Case-agreement theory.}

THAT IS, IF YOU WERE TO GET RID OF v YOU WOULD BE ELIMINATING BOTH THE ABILITY TO ASSIGN CASE TO AN OBJECT AND THE ABILITY TO ASSIGN A SUBJECT ROLE. THIS, INCIDENTALLY, IS ARGUABLY THE STRONGEST ARGUMENT THERE IS, INTERNAL TO ENGLISH, FOR DECOMPOSING VERBS, AT LEAST UP TO HALE & KAYSER'S EXTENT. WE RETURN TO THIS IN THE SEMINAR.

Each CFC also allows an extra SPEC beyond its s-selection: for C, a raised wh>-phrase; for T, the surface subject; for v>, the phrase raised by Object Shift (OS).

WHETHER CFCS ALLOW AN EXTRA SPEC, OR EXACTLY ONE EXTRA SPEC, IS AN IMPORTANT ISSUE TO KEEP IN MIND. OF COURSE, BARE PHRASE-STRUCTURE ALONE WOULD ALLOW FOR AS MANY SPECS AS YOU WANT, BUT PRESUMABLY THE SPECS HAVE TO BE JUSTIFIED BY SOME OTHER PROPERTY. FOR INSTANCE:

For T, the property is the Extended Projection Principle (EPP). By analogy, we can call the corresponding properties of C and v> EPP-features, determining positions not forced by the Projection Principle.

OF COURSE, NOT TO SAY THAT WE KNOW WHAT THAT MEANS--HENCE THE NAME 'PRINCIPLE' AND THE NOTION 'EXTENDED'. MORE PROFOUNDLY, AS I SAID ABOVE, WE DON'T EVEN REALLY KNOW WHY THE *PROJECTION* PRINCIPLE WORKS THE WAY IT DOES.

I will restrict attention to XP-positions, though a fuller picture might add X<0> as another case of EPP (see note 93).

DON'T BOTHER--TOO CRYPTIC TO UNDERSTAND AT THIS POINT.

EPP-features are uninterpretable (nonsemantic, hence the name), though the configuration they establish has effects for interpretation. Basic structural properties of CFCs are illustrated in the configuration (4), where H is the CFC, XP is the extra SPEC selected by its EPP-feature, and EA is the external argument selected by H = v>. Here we find the properties (5):

A RATHER CRYPTIC PASSAGE FOLLOWS, SO LET'S READ IT CAREFULLY.

(4) " = [XP [(EA) H YP]]

LET ME, FIRST, PUT THIS IN TREE NOTATION TO FOLLOW THE POINTS RAISED IN (5):

HP

/ \

XP H'

/ \

(EA) H'

/ \

H YP

(5)(i) If H is v>/C, XP is not introduced by pure Merge

IN OTHER WORDS, ONLY EXPLETIVE SUBJECTS ARE INTRODUCED IN T BY PURE MERGE. ACTUALLY, I WONDER WHAT CHOMSKY HAS IN MIND FOR whether.

(ii) In the configuration [<ß> T<ß>..."], ß minimal,

HERE WHAT WE WANT TO SEE IS WHAT ARE THE RELATIONS BETWEEN SOME T HEAD AND THE NEXT DOMAINS DOWN THE TREE--WHICH WILL BE VERY IMPORTANT IN THIS PAPER.

(a) if H is C, T<ß> is independent of "

THIS BASICALLY STATES THE TEMPORAL INDEPENDENCE BETWEEN ONE CLAUSE AND THE NEXT, TOGETHER WITH OTHER IMPORTANT ISSUES CONCERNING THE ISLANDHOOD OF THE EMBEDDED CLAUSE, AND SO ON. KEEP IN MIND A COUPLE OF SIMPLE FACTS: CPS EMBEDDED UNDER T ARE DOMAINS OF THETA AND CASE ASSIGNMENT, BINDING, (POSSIBLY) CONTROL, AND A VARIETY OF PHENOMENA RELATED TO CLITICS AND SIMILAR MORPHOLOGICAL DEPENDENCIES, AS WELL AS IMPORTANT LIMITATIONS ON A-MOVEMENT (OBVIOUSLY) AND PARTLY ALSO A'-MOVEMENT (REQUIREMENTS FOR SUCCESSIVE CYCLICITY, OR SUBJACENCY). CHOMSKY'S (iia) ESTABLISHES THIS INDEPENDENCE OF EMBEDDED CLAUSES, WHICH I WILL CALL 'PARA' (FOR 'PARATAXIS'). NEEDLESS TO SAY, NOBODY HAS A CLUE AS TO WHY PARA SHOULD OBTAIN WHERE IT DOES.

(b) if H is v>, T<ß> agrees with EA, which may raise to SPEC-T<ß> though XP cannot

THE SITUATION HERE IS THE CORE EPP, RELATING T TO THE NEXT v DOWN. WE'LL SEE LATER ON WHY XP CANNOT RAISE TO T, AND THIS ELEMENT ONLY AGREES WITH EA. NOTE THAT INTRODUCTION OF THE NOTION 'AGREE' HERE ALLOWS FOR A RELATION BETWEEN T AND v WITHOUT INVOLVING

DISPLACEMENT, WHICH IS EMPIRICALLY NECESSARY IN SOME LANGUAGES. FORMALLY, THE IDEA IS NO DIFFERENT FROM WHAT CHAPTER 4 CALLED THE ATTRACTION OF A FEATURE IN v, AND IT IS CURIOUS HOW CLOSE IT COMES TO SOME OLDER IDEAS, FOR INSTANCE RAPOSO & URIAGEREKA'S 1990 NOTION OF 'LONG-DISTANCE CASE ASSIGNMENT,' WHICH IN EFFECT WAS A FORM OF AGREEMENT UNDER EXCEPTIONAL GOVERNMENT. SO FAR AS I CAN SEE THE DOMAINS OF AGREE, FEATURE ATTRACTION, AND LONG-DISTANCE ASSIGNMENT ARE IDENTICAL, THE ONLY DIFFERENCE BEING IN TERMS OF WHAT ASPECT OF THE RELATION THE THEORY CONCENTRATES ON:

AGREE: ...X...[...Y...]... CHOMSKY 98

|_ _ _ _|

MATCHING REQUIREMENT

ATTRACT: ...X...[...Y...]... CHOMSKY 95

^_ _ _ _/

FEATURE DISPLACEMENT

ASSIGN: ...X...[...Y...]... RAPOSO & URIAGEREKA 90

_ _ _ _^

FEATURE LICENSING

IN ALL THREE OF THESE INSTANCES YOU NEED TO QUALIFY WHAT THE BRACKETS [...] ARE THAT (DIS)ALLOW THE RELATION--IN OTHER WORDS, SOMETHING ABOUT LOCALITY. THIS OF COURSE IS INDEPENDENT, AND ALL THREE VERSIONS OF THE SYSTEM COULD IN PRINCIPLE AGREE ON THAT. THIS IS ALL TO SAY THAT WHAT WOULD ALLOW YOU TO DECIDE AMONG THESE ALTERNATIVES IS SUBTLE, PERHAPS EVEN ONLY TECHNICAL. WE'LL SEE.

(c) if H is T<def>, XP raises to SPEC-T<ß> if there is no closer candidate ' for raising

THIS SITUATION REFERS TO 'DEFECTIVE' T, OR THE SORT OF T THAT APPEARS IN RAISING CONSTRUCTIONS, WHICH IS THUS INCAPABLE OF LICENSING NOMINATIVE CASE. AGAIN, WE'LL HAVE TO SEE WHY IN THOSE INSTANCES XP (WHICH CAN BE SEVERAL THINGS, DEPENDING ON EXACTLY WHAT IS THE COMPLEMENT OF THIS Tdef) CAN END UP RAISING ALL THE WAY UP, AS USUAL WHEN NO CLOSER ELEMENT INTERVENES, A MATTER THAT THE NEXT FOOTNOTE, I THINK, INVITES US NOT TO CONSIDER NOW:

{Irrelevant to our concerns here, for C (i) might be parametrically contingent on other operations (e.g., partial wh>-raising as in German). Other possible parametrization is put aside here.}

I DON'T KNOW WHAT (i) REFERS TO, AND HAVE A HARD TIME PARSING THIS FOOTNOTE.

Pure> Merge is Merge that is not part of Move. The relevant properties of T<ß> have to do with Case/agreement and EPP. In (iib), if EA does not raise, SPEC-T<ß> is introduced by pure Merge to satisfy EPP.

THAT IS, OF COURSE, THE SITUATION WITH OVERT EXPLETIVES.

The case of H = nondefective T is omitted in (ii): if (iia) holds for C, it holds for (nondefective) T selected by C.

THAT IS, WE DON'T CONSIDER NON-DEFECTIVE T IN (ii) BECAUSE:

WHEN (iia) HOLDS FOR C: IT ALSO HOLDS FOR LOWER (REGULAR) T:

[T... [...C...] ...] [T... [...C [...T...] ...] ...]

TP CP TP CP TP

INDEPENDENT INDEPENDENT

In fuller generality, ß in (5ii) should be taken to be the minimal ß containing " headed by any CFC H<ß>, which would therefore be either T or v>.

IN OTHER WORDS, WE ARE EXPLORING THE MINIMAL ß THAT CONTAINS " = [XP [(EA) H YP]], SUCH THAT ß IS HEADED BY ANY OF THE CFCS. GIVEN THE (AS OF YET NOT UNDERSTOOD) RESTRICTIONS ON SELECTION THAT WE ALL ASSUME, THE POSSIBILITIES ARE:

FOR " = C: FOR " = T: FOR " = v':

A. v' B. C' C. T'

/\ /\ /\

v CP CP TP T vP

NORMAL EMBEDDING NORMAL DECLARATIVES NORMAL TRANSITIVES

D. vP

/\

v TP

RAISING CONTEXTS

WHY CHOMSKY SAYS THAT THEREFORE THE HEAD OF ß MUST BE EITHER T OR v

IS UNCLEAR TO ME, SINCE IN B ABOVE IT IS CLEARLY C; I SUPPOSE CHOMSKY DOES NOT CONSIDER THIS INSTANCE BECAUSE, AS WE SAW IMMEDIATELY BEFORE, 'the case of H [i.e. the head of "] = nondefective T' IS BEING OMITTED, FOR THE REASONS POINTED OUT.

The relations of T<ß> to " extend partially to v> as well; specifically, ECM in (iic) as compared to raising to SPEC-T. We put this aside temporarily, for ease of exposition and because of some differences to which we return.

ALRIGHT, SO NOW LET'S SEE HOW THE PROPERTIES IN (5) ARE ACCOUNTED FOR.

Property (5i) follows in part from the theta-theoretic principle (6), which is implicit in the conception of theta roles as a relation between two syntactic objects, a configuration and an expression selected by its head

WE SEE PRINCIPLE (6) IN A MOMENT; NOTE THAT, FOR BETTER OR FOR WORSE, CHOMSKY IS COMMITTING TO A THEORY ALONG THE LINES OF HALE-KEYSER, CONTRA SEVERAL RECENT PROPOSALS THAT TREAT ROLES AS FEATURES (SEE THE NEXT FOOTNOTE). SINCE THAT WILL BE A CENTRAL TOPIC OF THIS SEMINAR, I WILL POSTPONE DISCUSSING IT FOR NOW.

{In effect, the Hale-Keyser theory adapted in MP> (chap 4.6). One consequence of this conception of Theta Theory is that the theta-criterion cannot be satisfied by raising an argument to a theta-position or by raising of "theta-features" (the existence of such features aside, I will suggest below that feature movement may not be possible). Other conceptions reject these conclusions, e.g., Hornstein (1997), Manzini and Roussou (1997).}:

(6) Pure merge in theta position is required of (and restricted to) arguments

I DO WANT TO POINT OUT THAT (6), COUPLED WITH THE FACT THAT NUMERATIONS ARE TAKEN TO BE SETS OF LEXICAL TOKENS AND FURTHERMORE OF *ALL* THE TOKENS IN A DERIVATION ARE TO BE FOUND IN THE NUMERATION (WHAT YOU MAY THINK OF AS A 'CONSERVATION' PROPERTY OF DERIVATIONS STATED IN TERMS OF NUMERATION/LF PAIRS), BRINGS BACK MUCH OF THE SUBSTANTIVE WORK OF D-STRUCTURE, NOT AS A SINGLE, ARTICULATED AND UNIFIED LEVEL OF REPRESENTATION, OF COURSE, BUT AS SOME SORT OF COMPONENT OF THE SYSTEM. IN PARTICULAR, KEEP IN MIND THAT (6) WAS TRUE BY DEFINITION OF D-STRUCTURE, AS WAS THE FACT THAT D-STRUCTURE WAS A COLLECTION OF LEXICAL TOKENS (OR YOU WOULDN'T HAVE PHRASE-MARKERS TO BEGIN WITH) AND THAT 'CONSERVATION' HELD TRIVIALLY, STRONGLY SO IF THE PROJECTION PRINCIPLE WAS ASSUMED. AT ANY RATE, YOU MUST ASK YOURSELF WHY (6) SHOULD HOLD. IT IS NOT AN OBVIOUS OPTIMALITY OF DESIGN PROPERTY (AT LEAST I DON'T KNOW OF ANY ARGUMENT TO THAT EFFECT), WHICH MUST THEN MEAN IT FOLLOWS FROM LEGIBILITY CONDITIONS. ONE IS OF COURSE TEMPTED TO THINK THAT (6) SHOULD FOLLOW FROM SUCH CONDITIONS ON LF, BUT IF LF IS A COLLECTION OF CHAINS (SETS OF PHRASE-MARKERS, AND THEREFORE BY DEFINITION NOT CONFIGURATIONAL) IT IS NOT ENTIRELY CLEAR HOW TO STATE SUCH A

LEGIBILITY CONDITION. WE RETURN TO THIS EXTENSIVELY.

We can therefore restrict attention in (5i) to expletives (EXPL); for arguments it follows from (6). (6) also applies to (5ii), restricting pure Merge in SPEC-T<ß> to EXPL.

THAT IS, OF COURSE, UNDER THE ASSUMPTION THAT T DOESN'T INVOLVE A ROLE (FOR MYSTERIOUS REASONS AT THIS POINT; NOTE YOU CANNOT JUST SAY THAT ONLY SUBSTANTIVE CATEGORIES INVOLVE ROLES, SINCE BY HYPOTHESIS v IS NOT SUBSTANTIVE AND YET IT INVOLVES A ROLE). NOTE ALSO THAT UNDER THIS VIEW ONE MIGHT ACCOMMODATE whether BASE-GENERATED AS MERGED TO C, ASSUMING THIS ELEMENT IS NOT AN ARGUMENT. LIKEWISE, PERHAPS ONE COULD KEEP TO RIZZI'S ANALYSIS OF why/how BASE-GENERATED IN THAT VERY POSITION, WHICH ACCOUNTS FOR SOME OF THEIR PUZZLING PROPERTIES. SIMILARLY FOR THE BASE-GENERATION OF WH-ELEMENTS ASSOCIATED TO SO-CALLED RESUMPTIVE PRONOUNS.

The property (5iia) is illustrated in (7), (iib) in (8), and (iic) in (9):

I WILL MOVE CHOMSKY'S TEXT AROUND HERE, TO ILLUSTRATE EACH OF THE EXAMPLES IN TURN WITH HIS OWN COMMENTS:

(7)(i) there are questions about [<"> what C [<TP> John read t>]]

(ii) there is a possibility [<"> that proofs will be discovered]

In (7), " = CP with H = C, and the relevant properties of matrix T (T<ß>) are independent of " in accord with (5iia). " is a closed system with regard to Case-agreement properties, determined internally without effect beyond.

AS I SAID, NOT JUST CASE-AGREEMENT, BUT ALSO THETA-RELATIONS, BINDING CONDITIONS, AND POSSIBLY OTHERS.

(8) T<ß> [<"> [<DO> the book] [[<SU> many students] [read t><DO>]]]

In (8), " is an OS construction with H = v> and DO = XP of (4)/(5iib). Under Holmberg's generalization V raises to T<ß>, which agrees with EA = SU. EA can then raise to SPEC-T<ß> ("many students read-pl the book (never)" or remain in situ with merge of expletive to satisfy EPP ("there read-pl the book (never) any students"), illustrating the two options for (5iib). The positions of DO, SU can be determined by left-edge adverbs.

{Translations of Icelandic examples based on Jonas (1996). It had been assumed that SU is higher than OB in these constructions, but Jonas found that the conclusions relied on improper choice of left-edge markers. This eliminates complications in MP> about the issue.}<JONAS = footnote>

I DON'T WANT TO GO INTO THIS, BUT I'VE HEARD REPEATEDLY THAT THE

ALLEGED FACTS ARE CONTROVERSIAL. LET'S ASSUME THEM IF ONLY TO SEE WHERE THEY LEAD.

(9)(i) T<ß>-is likely [<"> there to be a proof discovered]

(ii) T<ß> [{v>P} I expected [<"> there to be a proof discovered]]

In (9), " = TP with H = T<def> and XP = there>. XP raises to SPEC-T<ß> in (i), but not in (ii) with intervening ' = I>, yielding (10i,ii), respectively, which illustrate (5iic):

(10)(i) there is likely to be a proof discovered

(ii) I expected there to be a proof discovered.

Properties of matrix T (= T<ß>) depend on " in (9i) but not in (9ii), where it is the head of v>P (not T<ß>) that is related to ". The relation of T<ß> to " in (i) is analogous to that of v> to " in (ii), but there are differences to which we return. The order of object-subject in (8) follows from the assumption that Merge preempts the more complex operation Move.

REMEMBER THAT THIS ASSUMPTION COMES FROM CHAPTER 4, ALTHOUGH THERE THE WRONG REASON WAS GIVEN FOR WHY THE CONCLUSION SHOULD HOLD (THE IDEA THERE WAS THAT MERGE IS COSTLESS BECAUSE IT IS NEEDED FOR CONVERGENCE; THAT OF COURSE CONFUSES A TYPE OF OPERATION WITH A TOKEN OF ITS APPLICATION: EVEN IF THE TYPE OF OPERATION IS REQUIRED BY CONVERGENCE, IT IS NOT CLEAR WHY A TOKEN OF ITS APPLICATION SHOULD HAVE NO COST WHATSOEVER). THE SUGGESTION HERE IS RATHER DIFFERENT: MOVE INVOLVES SEVERAL SUB-OPERATIONS, INCLUDING MERGE; THEREFORE MOVE IS MORE COMPLEX, AND THUS ALL OTHER THINGS BEING EQUAL THE GRAMMAR PREFERS TO MERGE THAN TO MOVE.

When the derivation has reached the stage (11), two operations have to take place for an OS construction, each creating SPEC, Merge of the external argument EA and raising of object DO (Move):

(11) [v> [V DO]]

If Merge applies first, the order is as in (8ii). The examples of (9)/(10) contrast with (12):

(12)

(i) *there is likely [<"> a proof to be discovered]

(ii) *I expected [<"> t> to be a proof discovered]

(iii) I expected [<"> a proof to be discovered]

Suppose the derivation has reached reached [SIC] the stage (13), analogous to (11), with T<"> = T<def>:

(13) [T<"> [be a proof discovered]]

EPP requires that something occupy SPEC-T<">. Two options are available: merge there> or move a proof>. Preference of Merge over Move selects the former. Accordingly, (9i)/(10i) is permitted and (12i) is barred. But Merge of an argument in SPEC-T<"> violates the theta-theoretic condition (6). Therefore (12ii) is barred. Either an expletive is merged yielding (10ii), or Move applies yielding (12iii). The choice depends on whether or not an expletive is available in the initial lexical array -- the first time step (III) of (3) enters.

THAT IS, THIS IS THE FIRST TIME HAVING A NUMERATION HAS BOUGHT US SOMETHING--THE ARGUMENT IS FAMILIAR FROM CHAPTER 4. NOTICE, WITHOUT THE NOTION OF A LEXICAL ARRAY (BE IT A NUMERATION, I.E. COLLECTION OF TOKENS, OR SOMETHING LESS SPECIFIC), IT WOULD BE INCOHERENT TO SAY 'whether or not an expletive is available in the initial lexical array'. AS A RESULT WE WOULD FALL INTO A PROBLEM WHICH, IN PAST SEMINARS, WAS NOTED BY JUAN CARLOS CASTILLO: WHY SHOULD YOU EVER MOVE IF YOU COULD SELECT AN EXPLETIVE, THUS RESULTING IN A CHEAPER DERIVATION? OF COURSE, IF YOU HAVE A LEXICAL ARRAY THE ISSUE DOESN'T ARISE: IF YOU HAVE THE EXPLETIVE YOU DON'T MOVE, IF YOU DON'T HAVE THE EXPLETIVE YOU HAVE TO MOVE.

Four kinds of complexity considerations enter into this account:

(14)

(i) Simple operations preempt more complex ones

THIS ONE IS TRIVIAL. THE NEXT ONE IS STRAIGHTFORWARD, AND CAN BE SEEN IN VARIOUS WAYS (E.G. THE MINIMAL LINK CONDITION PREVENTING OPERATIONS OVER CLOSER ITEMS):

(ii) Search space is limited (locality)

(iii) Access to the feature set [F] is restricted by (3)

THAT ONE PROVIDES A RATIONALE FOR THE LEXICAL ARRAY, AS WE SAW. BUT OF COURSE THE MOST IMPORTANT ONE, GIVEN THE CONCEPTUAL POINTS RAISED IN THIS PAPER, IS THE FOLLOWING:

(iv) Computation is locally determined (no look-ahead)

The conclusion (iv) follows from (i-iii) and (6), all plausible principles of some generality; in the background is the derivational approach (cyclicity).

MORE THAN IN THE BACKGROUND, AS YOU'LL SEE SHORTLY.

Alternative analyses have been proposed, but to my knowledge they are not locally determinable or introduce special or dubious assumptions.

HERE I'D LIKE TO SAY SOMETHING ELSE, WHICH CHOMSKY AS FAR AS I KNOW HASN'T EVER SAID--AND WOULD PROBABLY DISAGREE WITH. THE DERIVATIONAL APPROACH, WITNESSED AT ITS BEST IN THE CYCLIC PROPERTY, IS SO MUCH MORE INTERESTING! I DON'T MEAN JUST BECAUSE THE OPERATIONS ARE NEAT AND SO ON, BUT BECAUSE OF A KIND OF QUESTION THAT FEW PEOPLE WORRY ABOUT, BUT IN MY OPINION IS VERY INTRIGUING: WHAT KIND OF SENSE DOES *ANY* OF THIS MAKE FOR A NATURAL SYSTEM? AS THE COURSE PROCEEDS, I'LL TRY TO ARGUE THAT IT ACTUALLY MAKES A LOT OF SENSE IF THE SYSTEM IS DERIVATIONAL, GIVEN A DICTUM WHICH, AS FAR AS I KNOW, WAS FIRST COHERENTLY ARTICULATED BY D'ARCY THOMPSON AT THE TURN OF THE CENTURY: IN LIFE, ELEGANT FORM IS THE HISTORY OF A REGULAR PROCESS. THAT MIGHT SOUND DARWINIAN, BUT NOTHING IN THE CLASSIC DARWINIST APPROACH GIVES YOU WHAT YOU MAY CALL STRUCTURALLY ELEGANT FORM (AT BEST, YOU'D GET FUNCTIONALLY ADEQUATE FORMS, PERIOD--AT WORST, A HODGE-PODGE OF NOT-GOD-KNOWS-WHAT). WHAT THOMPSON HAD IN MIND WAS THE NATURAL HISTORY OF A PHYSICO-CHEMICAL PROCESS, OR SOMETHING OF THAT ILK. THERE ARE MANY SYSTEMS OUT THERE WHICH PLAUSIBLY FALL UNDER THAT CATEGORY (BODY PLANS, SCALING LAWS, REGULARITIES IN PHYSIOLOGY, AND--THE HOPE IS, THOUGH THIS HAS NEVER BEEN PROVED--REGULARITIES IN BEHAVIOR). FOR THOSE SYSTEMS, A COMPUTATIONAL/DERIVATIONAL TREATMENT IS STANDARD. USUALLY, THIS IS DONE FOR MODELLING PURPOSES. BUT RESEARCHERS ARE EXPLORING THE POSSIBILITY OF EXTENDING THE ACCOUNT TO ACTUAL DEVELOPMENTAL STORIES, AND SOME ARE EVEN PLAYING WITH EVOLUTIONARY SCENARIOS IN THOSE TERMS. OF COURSE, WE ARE FAR, FOR REASONS I MENTIONED ABOVE, FROM COMING CLOSE TO THIS. STILL, A FULLY DEVELOPED DERIVATIONAL SYSTEM FITS RIGHT IN. IF YOU ARE WORRIED ABOUT WHETHER THIS IS REALISTIC, IN THAT THE KIND OF COMPUTATIONAL/DERIVATIONAL SYSTEM WE'RE DEVELOPING WOULD NOT CORRESPOND TO WHATEVER GOES ON IN PERFORMANCE, THINK THREE THINGS. (I) THOSE OTHER COMPUTATIONAL MODELS THAT EXIST OUT THERE FOR VARIOUS FORMS DON'T NECESSARILY CORRESPOND DIRECTLY WITH WHAT HAPPENS 'NOW'; SOME DO (PROTEIN-FOLDING MODELS) WHILE OTHERS DON'T (FRACTAL SYSTEMS THAT PREDICT SCALING LAWS); IT IS AT ANY RATE PERFECTLY OKAY TO IDEALIZE IN SOMETHING AS ABSTRACT AS THE COMPETENCE SYSTEM--YOU'RE NOT WASTING YOUR TIME, AND YOU'LL MAP IT TO WHATEVER YOU NEED TO MAP IT WHEN TIME COMES (AND INCIDENTALLY, YOU MIGHT END UP MAPPING IT INTO SOMETHING AS ABSTRACT AS A DEVELOPMENTAL SYSTEM OR EVEN AN EVOLUTIONARY SYSTEM); (II) IN ANY CASE, THE MORE CYCLIC A MODEL OF COMPETENCE IS, THE CLOSER IT ACTUALLY IS TO A PERFORMANCE MODEL, PARTICULARLY IF (AS IT SEEMS) THE CYCLES CORRESPOND TO NATURAL PERFORMANCE UNITS. (III) WE REALLY HAVE NO IDEA OF WHAT'S ULTIMATELY GOING ON 'NOW'; TO HAVE A BIAS FOR A REPRESENTATIONAL SYSTEM BECAUSE, SOMEHOW, IT SEEMS TO 'NATURALLY CORRESPOND TO KNOWLEDGE THAT PERFORMANCE USES' IS A BASICALLY MEANINGLESS STATEMENT--INDEED, IF THE PHILOSOPHICAL WORRY ABOUT

REPRESENTATIONS HAS ANY BITE TO IT, THE OPPOSITE MAY WELL BE THE CASE (ALTHOUGH OF COURSE NOBODY KNOWS). END OF THE CONCEPTUAL PLUG FOR THE SYSTEM.

By the reasoning just reviewed, other constructions similar on the surface to (12iii) could be legitimate, for example (15), where V is a seem>-type verb with an optional SPEC

{The counterpart of the to>-phrase in "seems to-me [NP to...]." This phrase is sometimes described as the complement of V, but we assume it to be a SPEC in a Larsonian shell. See note <LIGHT>.}

that is raised to matrix SPEC in (i) and missing in (ii):

(15)

(i) me(DAT) V [<TP> several people to be in the room]

(ii) there V [<TP> several people to be in the room]

(15i) is similar to (12iii): Move applies in embedded TP in preference to Merge of me>, in accord with (6);

THIS IS A FUNNY USE OF THE WORD preference, SINCE AS FAR AS I CAN SEE (6) IS A CONVERGENCE CRITERION, AND THUS THE MERGE OF me IS DEMANDED BY THE SYSTEM--IS NOT A PREFERENCE AT ALL, IN THE TECHNICAL SENSE OF THE TERM.

then merges as SPEC of the seem>-construction and raises to matrix subject.

THIS WOULD BE A QUIRKY CASE STRUCTURE.

Case (ii) could arise if the embedded clause is a multiple-subject construction MSC, in which both Merge (of expletive) and Move (of several people>) have applied, the expletive then raising to matrix SPEC. Both possibilities appear to be legitimate in languages with the relevant properties.

{On (i), see Sigurdsson (1996). On (ii), see MP>, chap. 4.9 (168).

NOTE, THAT WOULD GIVE YOU SOMETHING LIKE there seem several people to be in the room, WHICH, THE CLAIM IS, IS GOOD IN ICELANDIC, ALTHOUGH NOTE: WITH A DIFFERENT STRUCTURE FROM THAT IN THE BAD (12i), SINCE IN THIS INSTANCE THE PLEONASTIC STARTS DOWN AND MOVES UP, WHEREAS IN THE BAD ENGLISH SENTENCE THE PLEONASTIC STARTS UP. WHY ENGLISH DOESN'T ALLOW THE ICELANDIC STRUCTURE WOULD RELATE TO WHY IT DISALLOWS, MORE GENERALLY, MULTIPLE SUBJECT STRUCTURES (THAT IS, TRANSITIVE EXPLETIVES).

The structure of subjects in MSCs raises a large range of questions put aside here.}

Control infinitivals, I have assumed, fall together with finite clauses, headed by C selecting

nondefective T (with tense-modal structure and a full complement of N-features).

THAT CONTROL INFINITIVALS INVOLVE SOMETHING AKIN TO A FINITE CLAUSE HAS BEEN ARGUED BY MARTIN (1996), FOLLOWING IDEAS THAT GO BACK TO STOWELL (1983). STILL, MUCH MORE HAS TO BE SAID, OF COURSE, ABOUT WHY THE RELEVANT TPS INVOLVE A DEFECTIVE CASE SYSTEM. WE SHOULD RETURN TO THIS.

Like other Cps, they generally undergo movement and clefting and can appear as root expressions (typically with wh-phrase> SPEC or as discourse fragments), and structural Case is assigned to the subject of T.

RELEVANT ITALIAN EXAMPLES HERE GO BACK TO RIZZI (1982), AND CAN BE TRANSLATED INTO ENGLISH AS FOLLOWS:

(I) A. ONLY TO DRINK HIS BEER DOES HE ULTIMATELY WANT!

B. *ONLY TO LIKE BEER DOES HE ULTIMATELY SEEM!

(II) A.?HE CLAIMED HE WANTED TO BUY A CAR, BUT TO BUY A CAR..., HE CERTAINLY DIDN'T WANT THAT, JUDGING FROM WHAT HE BOUGHT...

B.*HE CLAIMED HE WAS LIKELY TO LIKE ESCARGOTS, BUT TO LIKE ESCARGOTS... HE CERTAINLY WASN'T LIKELY THAT, AFTER WHAT HE DID TO AVOID EATING THEM...

(III) A. TO TALK TO MARY IS WHAT JOHN WANTS.

B. *TO LIKE BEER IS WHAT JOHN SEEMS.

(IV) A. (WHO) TO HIRE PROPERLY? I DON'T KNOW...

B. *(WHO) TO HIRE PROPERLY? I'M NOT LIKELY...

These properties are common to Cps and distinguish them from raising/ECM infinitivals headed by T<def>, lacking C and tense structure and assigning no Case to subject, and lacking the distributional freedom of CP.

{See MP>, Martin (1996). Some of these distinctions have been attributed to trace-government, but that mechanism is not available here.}

THAT IS, TRADITIONALLY YOU WOULD HAVE SAID THAT IF YOU MOVE AROUND THE PHRASE CONTAINING THE TRACE OF RAISING THEN YOU'LL VIOLATE THE ECP; BUT NOW YOU DON'T HAVE THE ECP, SO YOU NEED A DIFFERENT ACCOUNT. IN ANY CASE, WHAT'S IMPORTANT HERE IS THAT THE PHENOMENOLOGY OF CONTROL AND RAISING IS KNOWN TO BE DIFFERENT, SO YOU WANT YOUR

THEORY TO CAPTURE THIS, REGARDLESS OF HOW YOU ULTIMATELY DO IT.

We also have such convergent constructions as (16), analogous to (7) and contrasting with (12i):

I'LL REPEAT (7) AND (12I) FOR COMPARISON:

(7)

(i) there are questions about [<"> what C [<TP> John read t>]]

(ii) there is a possibility [<"> that proofs will be discovered]

(12)

(i) *there is likely [<"> a proof to be discovered]

(ii) *I expected [<"> t> to be a proof discovered]

(iii) I expected [<"> a proof to be discovered]

(16)

(i) it's fun [<"> PRO to [t> go to the beach]]

(ii) it's about time [<"> PRO to [leave t>]]

(iii) it was decided [<"> PRO to be executed t> at dawn]

THIS IS A CENTRAL DOMAIN IN THIS PAPER, WHAT LATER ON WILL BE CALLED A 'PHASE' IN THE DERIVATIONAL SYSTEM. FOR NOW THE POINT CHOMSKY IS TRYING TO ESTABLISH IS THAT CONTROL STRUCTURES BEHAVE PHASE-LIKE, HENCE ANALOGOUS TO (7), AND UNLIKE (12).

Raising is possible throughout in the closed system " (as in (7)). In (iii), at least, PRO is controlled by an implicit argument: it can mean that the prisoners decided that they would be executed at dawn, but not that we decided that they would be.

EVEN IF WE AGREE WITH THE JUDGEMENT HERE, IT IS NOT COMPLETELY CLEAR TO ME WHY CHOMSKY EMPHASIZES THIS FACT, UNLESS HE'S IMPLICITLY ADDRESSING THEORIES OF PRO OF THE SORT HE ALLUDED TO IN FN. 35.

A problem throughout the whole account is why raising is ever> possible, if Agree and Merge preempt Move.

NOTE, THE SAME LOGIC THAT WOULD HAVE MERGE PREEMPT MOVE WOULD EXTEND TO AGREE, AS DID FOR ATTRACT VS. MOVE IN CHAPTER 4 (YIELDING PROCRASTINATE).

The question is answered in part by the theta-theoretic principle (6), which bars pure merge of arguments in non-theta positions, and correspondingly restricts Move to such positions.

SO AGAIN, (6) IS SEEN AS A CONVERGENCE CONDITION, MAKING PREFERENCES

IRRELEVANT.

Choice of Move over Agree follows from presence of EPP-features, where pure Merge is inapplicable.

THAT IS, IN SITUATIONS WHERE EPP DEMANDS A CATEGORY (FOR WHATEVER REASON) AND IT IS IMPOSSIBLE TO MERGE, THE SYSTEM WILL NOT BE ABLE TO INVOKE AGREE (WHICH DOESN'T INVOLVE CATEGORIES) OR MERGE (WHICH REQUIRES A ROLE IF THE CATEGORY IS OF THE RELEVANT SORT) AND WILL THUS BE FORCED TO MOVE. AS FAR AS I CAN SEE, THIS WILL ONLY ARISE FOR v (AND IT MIGHT FOR C DEPENDING ON OTHER ASSUMPTIONS).

The remaining question, then, is why Merge of EXPL does not always bar Move. That question is partly answered by the initial choice of lexical array: it may or may not make EXPL available. But that cannot be the whole story, as illustrated in (7) and (16), where EXPL is available in the lexical array but Move takes place in the embedded phrase

THIS IS, AS FAR AS I KNOW, AN INTERESTING PROBLEM THAT WAS FIRST NOTED IN THE CONTEXT OF CHOMSKY'S CLASS A COUPLE OF YEARS AGO, INDEPENDENTLY BY ALEC MARANTZ AND JUAN ROMERO.

{Observe that (9), (10), (12iii) appear problematic because of raising of "a proof" from DO of "discover." There are reasons to suppose that the actual structure at this stage is the expected "there to be [discovered a proof]" as in similar languages, with the preferred English construction "a proof [discovered t>]" formed outside the system we are now considering. I will assume so, leaving the issue to the side here.}<PROB = footnote>

THIS FOOTNOTE SIDESTEPS A VERY IMPORTANT TOPIC THAT I'D LIKE TO COME BACK TO. NOTE, OF COURSE, THAT IF a proof to be discovered t IS THE ACTUAL STRUCTURE, THEN WE HAVE A RATHER SERIOUS PROBLEM TO DEAL WITH, SINCE IN FACT MOVEMENT OF a proof HAS TRUMPED MERGE. I WILL GO ALONG WITH CHOMSKY FOR NOW, ALTHOUGH I THINK THE OTHER LANGUAGES HE'S TALKING ABOUT (ICELANDIC, ROMANCE) DIFFER FROM ENGLISH IN SIGNIFICANT RESPECTS, AND IN FACT THE STRUCTURE IN QUESTION IS, IN ENGLISH, AS YOU SEE IT. BUT I'LL RETURN TO THIS LATER ON IN THE SEMESTER, WHEN I TRY TO GIVE AN ACCOUNT IN A SYSTEM THAT IS EVEN MORE DERIVATIONAL THAN CHOMSKY'S.

A straightforward solution would be to take the derivational approach still more seriously and extend further the procedures (3) that reduce access to the domain of L. Suppose we select LA as before, under (III) of (3); the computation need no longer access the lexicon. Suppose further that at each stage of the derivation a subset LA<{i}> is extracted, placed in active memory (the "work space"), and submitted to the procedure L.

THIS IS THE KEY--A CYCLIC ACCESS TO THE NUMERATION. NOTE, INCIDENTALLY, THAT WE HAVE JUST AGREED TO PROVIDE THE NUMERATION WITH ONE MORE PROPERTY. IS IT SUBSTANTIVE OR IS IT FORMAL, AND IF THE LATTER, DOES IT FOLLOW FROM DESIGN SPECIFICATIONS? LET'S GO ON:

When LA<{i}> is exhausted, the computation may proceed if possible. Or it may return to LA and extract LA<{j}>, proceeding as before. The process continues until it terminates.

NOTE: THE FACT THE SYSTEM KNOWS WHEN TO TERMINATE IS IN PART A REFLEX OF HAVING A NUMERATION--THEREFORE AS I SAID BEFORE ANOTHER IMPORTANT PROPERTY OF THE LEXICAL ARRAY.

Operative complexity in some natural sense is reduced, with each stage of the derivation accessing only part of LA.

THIS IS MEANT AS THE EXPLANATION (OR AT LEAST THE MOTIVATION) FOR WHY WE HAVE A CYCLIC ACCESS TO THE NUMERATION. OF COURSE, THERE MAY WELL BE SEVERAL OTHER WAYS OF REDUCING OPERATIVE COMPLEXITY. AS A MATTER OF FACT, IF THE SYSTEM REDUCES OPERATIVE COMPLEXITY AS DRASTICALLY AS IT IS NOW IMPLIED (AND THIS WILL BECOME EVEN MORE DRASTIC BELOW), IT IS NOT ALTOGETHER CLEAR WHY THE SYSTEM NEEDS THE, AS IT WERE, 'INTERMEDIATE' STEPS OF REDUCING COMPLEXITY. FOR INSTANCE, WHY HAVE A NUMERATION? OF COURSE, THE REASON IS EMPIRICAL (BASICALLY, CASTILLO'S OBSERVATION), BUT WE STILL NEED TO MOTIVATE, AS MINIMALISTS, WHY THINGS ARE THE WAY THEY HAVE TURNED OUT TO BE.

If the subarray in active memory does not contain EXPL, then Move can take place in the corresponding stage; if it does, Merge of EXPL preempts Move.

THAT'S OF COURSE THE EXPLANATION OF THE MARANTZ/ROMERO PUZZLE. BUT CRUCIALLY WE NEED TO SAY, NOW, WHAT COUNTS AS A VALID CYCLIC ACCESS TO THE NUMERATION, OR THE DISTINCTION BETWEEN THE EXAMPLES IN (7) AND (12) WILL COLLAPSE.

{Why not dispense with LA, just selecting subarrays cyclicly?

THIS IS THE ISSUE JUST POSED.

Apart from the general considerations about access reduction already discussed,

THOSE, AS WE SAW, ARE NOT ENTIRELY CONVINCING IN THE ABSENCE OF FURTHER THEORETICAL APPARATUS.

there is a more specific reason: chain properties can be reduced in significant part to identity if lexical arrays are enriched to numerations.

THIS IS VERY IMPORTANT, AND THE INTUITION IS PROBABLY THE RIGHT ONE IF ONE'S SYSTEM HAS CHAINS. THE ISSUE IS THIS. YOU NEED SOME KIND OF CONSERVATION LAW THAT TELLS YOU SOMETHING TO THE EFFECT THAT WHATEVER CHAINS YOU HAVE AT LF (AND ONLY THOSE) ARE THE RESULT OF MANIPULATING WHATEVER LEXICAL FEATURES YOU HAD IN YOUR NUMERATION. WITH THAT IDEA IN MIND, CHAIN IDENTIFICATION REDUCES TO SOME KIND OF INPUT/OUTPUT CONDITION ON DERIVATIONS PREVENTING THEM FROM 'RUNNING WILD', ARBITRARILY CREATING OR DESTROYING ALL SORTS OF

STUFF (WHAT I'M CALLING A CONSERVATION LAW, WHICH IS VERY REASONABLE FOR A NATURAL SYSTEM TO HAVE--E.G. THE FIRST LAW OF THERMODYNAMICS). IF YOU HAVE A NUMERATION (OR D-STRUCTURE, FOR THAT MATTER) THIS CHAIN IDENTIFICATION PROCEDURE IS ENTIRELY TRIVIAL, PARTICULARLY IF, AS WE SEE BELOW, YOU REDUCE THE SPACE WHERE CHAINS CAN EXIST RATHER DRAMATICALLY IN TERMS OF PHASES AND SIMILAR CONSTRUCTS. OTHERWISE, IT IS MUCH HARDER TO SEE HOW A CHAIN COMES ABOUT, AND YOU NEED A SEPARATE COMPONENT OF THE SYSTEM DEVOTED TO DOING PRECISELY THAT. FOR EXAMPLE, BRODY'S OR KOSTER'S REPRESENTATIONAL ALGORITHMS, OR NUNES'S (1995) DERIVATIONAL PROCEDURE. I WILL HAVE LITTLE TO SAY ABOUT THE REPRESENTATIONAL ALGORITHMS, BASICALLY BECAUSE I'M TRYING TO EXPLORE A DIFFERENT KIND OF SYSTEM. AS FOR NUNES'S ALTERNATIVE, THE PROBLEM WITH IT IS THAT IT VIOLATES THE ASSUMPTION OF 'COMPLETENESS' IN THE MINIMALIST PROGRAM. IN EFFECT, NUNES CREATES A SEPARATE LEVEL OF REPRESENTATION FOR THE IDENTIFICATION OF CHAINS (UNDER CERTAIN REPRESENTATIONAL CONDITIONS, SUCH AS COMMAND AND IDENTITY OF TOKENS). HOWEVER, SUCH A LEVEL WOULD HAVE TO EXIST, BY DEFINITION, BETWEEN THE LAST SYNTACTIC LEVEL AND THE FIRST SEMANTIC REPRESENTATION, ASSUMING THE LATTER IS BUILT FROM CHAINS. THEREFORE THE CHAIN LEVEL IS THE INTERFACE LEVEL, AND THEREFORE WHAT WE USUALLY CALL LF (WHERE PROCEDURES OTHER THAN CHAIN IDENTIFICATION TAKE PLACE) WOULD NOT BE THE INTERFACE, IN VIOLATION OF COMPLETENESS. THE ONLY WAY AROUND THIS, WITHOUT ASSUMING SOMETHING LIKE A NUMERATION, WOULD BE TO SAY THAT PROCEDURE L INTERFACES NOT JUST AN INTENTIONAL/CONCEPTUAL SYSTEM, BUT ALSO, CRUCIALLY, SOME OTHER SYSTEM WHERE CHAIN IDENTIFICATION IS REQUIRED. PERHAPS VALDUVI'S 'INFORMATION STRUCTURE' IS ONE SUCH LEVEL, IF CHAINS SOMEHOW HAVE TO DO WITH THEME/RHEME RELATIONS. WHILE THAT MAY WELL BE POSSIBLE FOR EPP FEATURES (MORE ON THAT LATER ON), I DOUBT THAT IT WOULD BE HE GENERAL TREATMENT FOR UNINTERPRETABLE FEATURE CHAINS.

To achieve the same result with cyclic choice of successive subarrays requires continual access to the full lexicon and memory of how many times each item has been selected.}

I THINK THE DIFFICULTIES I'VE POSED ARE MORE SERIOUS THAN SOMETHING HAVING TO DO WITH MEMORY.

The next step is to characterize the subarrays LA<{i}> that can be selected for active memory. LA<{i}> should determine a natural syntactic object SO. Perhaps the simplest and most principled choice is to take SO to be the closest syntactic counterpart to a proposition:

I SUSPECT THIS, EVEN IF ULTIMATELY RIGHT ON EMPIRICAL GROUNDS, IS WRONG FOR TWO REASONS: (I) THE CHOICE FOLLOWS FROM NOTHING WITH A DESIGN PROPERTY (REMEMBER, WE'RE DECIDING HOW BIG THE CYCLIC ACCESS TO THE NUMERATION SHOULD BE). AND (II) IN ANY CASE, THE NOTION proposition IS NEITHER SYNTACTIC NOR, IF TRUTH BE TOLD, PARTICULARLY WELL DEFINED

(ONE OF THOSE CONCEPTS WE HAVE INHERITED FROM TRADITION). (IF YOU'RE THINKING A PROPOSITION IS JUST A PREDICATE AND ITS ARGUMENT JUDGED TRUE OR FALSE, THINK ABOUT EACH OF THOSE PRIMITIVES: WHAT DOES 'PREDICATE' ULTIMATELY MEAN? TRUTH? ETC.) HERE THE CONCEPT WILL BE CONFOUNDED EVEN FURTHER, GIVEN THIS:

either a verb phrase in which all theta roles are assigned or a full clause including tense and force. Call these objects "propositional."

OF COURSE, WE CAN *CALL* THEM WHAT WE WANT, BUT THAT DOESN'T MAKE THEM PROPOSITIONAL. I FOR ONE HAVE A HARD TIME UNDERSTANDING WHY A VERB-PHRASE IS MORE OR LESS PROPOSITIONAL THAN A TENSELESS CLAUSE. INDEED, IT WOULD SEEM TO ME THAT IN

(i) Everyone doesn't seem likely to win the lottery.

THERE IS A READING NATURALLY PARAPHRASABLE AS:

(ii) It doesn't seem likely that everyone will win the lottery.

WHERE, IN LOGICAL FORM AT LEAST (WHICH IS WHERE PROPOSITIONS MATTER) WE HAVE A PROPOSITION OF PRECISELY THE SAME SORT AS WE DO CORRESPONDING TO THE EMBEDDED CLAUSE IN (II). ANYWAY, FOR WHAT IT'S WORTH, I'LL TRY TO PERSUADE YOU THAT THERE IS A BETTER WAY OF DEFINING THE CYCLIC ACCESS TO THE NUMERATION, BUT I WILL DO THAT LATER ON IN THE SEMESTER. IN ANY CASE, IT'S CLEAR THAT WE *WANT* SOMETHING LIKE CHOMSKY'S T AND POSSIBLY ALSO v--THAT'S NOT WHAT WORRIES ME; THE QUESTION IS HOW WE GET THAT RESULT TO FOLLOW.

LA<{i}> can then be selected straightforwardly: LA<{i}> contains an occurrence of C or of v>, determining clause or verb phrase -- exactly one occurrence if it is restricted as narrowly as possible, in accordance with the guiding intuitions. Take a phase> of a derivation to be a syntactic object SO

CAREFUL WITH THE TERMINOLOGY. SYNTACTIC OBJECT MEANT SOMETHING TOTALLY DIFFERENT IN CHAPTER 4.

derived in this way by choice of LA<{i}>. A phase is CP or v>P, but not TP or a verbal phrase headed by H lacking N-features and therefore not entering into Case/agreement checking: neither finite TP nor unaccusative/passive verbal phrase is a phase.

IN ESSENCE, THAT'S HOW WE'D LIKE TO UNDERSTAND A PHASE: AS THE MINIMAL DOMAIN THAT HAS A COMPLETE SET OF RELEVANT FEATURES. THE ISSUE OF PROPOSITIONALITY IS REALLY BESIDE THE POINT, AS FAR AS I CAN SEE. TO BE HONEST, THOUGH, EVEN IF WE SUCCEED IN REFINING THE DEFINITION, SOMEONE COULD AND SHOULD ASK *WHY* THOSE FEATURES ARE CHECKED IN PRECISELY THE CATEGORIES THAT THEY ARE CHECKED, AND NOT OTHERS--THAT'S A TOUGH ONE, AND WE SHOULD RETURN TO IT.

Suppose phases satisfy a still stronger cyclicity condition:

> (17) The head of a phase is "inert" after the phase is completed, triggering no further operations.

THIS IS VERY MUCH IN THE SPIRIT OF THE MULTIPLE SPELL-OUT (MSO) SYSTEM SKETCHED IN URIAGEREKA (1995/FORTHCOMING) (HENCEFORTH ALSO MSO), ALTHOUGH THERE ONE THING IS DIFFERENT: A *MOTIVATION* IS GIVEN FOR WHY PHASES SHOULD BE "INERT": ASSUMING THEY SPELL-OUT, THE STRUCTURE IS LITERALLY GONE FROM SYNTACTIC COMPUTATION. THE REASON FOR ASSUMING THAT PHASES SPELL-OUT IS TO DERIVE KAYNE'S LCA--WHICH CHOMSKY ISN'T OBVIOUSLY ASSUMING ANY LONGER (FOR REASONS THAT WE WILL RETURN TO). OF COURSE, IF THE ENTIRE PHASE SPELLS OUT, THEN THE ENTIRE PHASE IS GONE FROM COMPUTATION. FOR CHOMSKY, IN CONTRAST, IT IS ONLY A CHUNK OF THE PHASE THAT IS INERT. YOU WILL SEE THAT CHOMSKY CANNOT GET THE ENTIRE PHASE TO BE INERT BECAUSE HE WANTS TO ALLOW SUCCESSIVE-CYCLIC MOVEMENTS OUT OF PHASES. IN THE MSO SYSTEM, SUCCESSIVE CYCLICITY WAS ACHIEVED THROUGH SIDEWARD (S) MOVEMENT, SOMETHING THAT I WILL RETURN TO LATER ON IN THE SEMESTER. AT ANY RATE, S MOVEMENT IS INDEPENDENT OF MSO, SO THAT ONE COULD BE RIGHT WHILE THE OTHER IS WRONG; SIMILARLY, CHOMSKY'S SPECIFIC PHASES ARE COMPATIBLE WITH S MOVEMENT, WHICH ADDS AN INTERESTING TWIST TO IT ALL. IT SHOULD BE SAID, ALSO, THAT NOTHING IN THE LOGIC OF MSO FORCES JUST NON-COMPLEMENTS TO SPELL-OUT. INDEED, IN URIAGEREKA (1998)--THE NLLT ARTICLE ON BASQUE SPECIFIERS--IT WAS SHOWN THAT ENTIRE CLAUSES CAN SPELL-OUT PARTIALLY IN CERTAIN CIRCUMSTANCES THAT WE RETURN TO. SIMILARLY, NUNES AND URIAGEREKA (FORTHCOMING) (HENCEFORTH N&U)--AN ARTICLE FOR SYNTAX ON PARASITIC GAPS--SHOW THAT COMPLEMENTS CAN SPELL-OUT PARTIALLY IF THEY ARE GOING TO MOVE, IN ORDER TO ACHIEVE UNIFORM CHAINS. DIFFERENTLY PUT, ONE THING IS TO HAVE AN MSO SYSTEM AND A DIFFERENT ONE IS TO ASK WHAT ARE THE SPELL-OUT DOMAINS; IF THE LOGIC OF THE ORIGINAL PAPER IS RIGHT, THESE DOMAINS (CALLED 'CASCADES' IN THAT CONTEXT) SHOULD LITERALLY *EMERGE* AS A RESULT OF VARIOUS DERIVATIONAL DYNAMICS--THEY SHOULDN'T HAVE TO BE DEFINED. THAT'S OF COURSE DESIRABLE IN A MINIMALIST SYSTEM. ONE OTHER THING, WHAT CHOMSKY CALLS A 'PHASE' (DEFINED IN TERMS OF CYCLIC ACCESS TO A NUMERATION) WAS CALLED A 'FLOW' IN THE EARLY MSO MANUSCRIPT. I LIKE CHOMSKY'S TERM BETTER, SO I'LL USE 'PHASE' THE WAY HE DOES. AT THE SAME TIME, I'LL USE THE TERM 'CASCADE' TO REFER TO A CHUNK OF STRUCTURE, POSSIBLY SMALLER THAN A PHASE BUT NEVER LARGER, WHICH UNDERGOES MSO FOR SOME REASON OR OTHER. AN INTERESTING, AND DIFFICULT EMPIRICAL ISSUE IS WHETHER PHASES COINCIDE WITH CASCADES.

A phase head cannot trigger Merge or Attract in a later phase,

I SUSPECT THIS USE OF THE WORD Attract IS JUST A RESIDUE FROM THE PAST...

and we can restrict attention to phases in which all selectional requirements are satisfied,

including EPP for T (by virtue of cyclicity) and for v>/C, and selection of external argument for v> if required; otherwise the derivation crashes at the phase level.

AGAIN, CHOMSKY INSISTS (AND NOW YOU CAN SEE WHY) ON THE IDEA THAT YOU CAN CRASH at the phase level. MIND YOU, THIS level IS NO LEVEL OF REPRESENTATION, AND IN CHOMSKY'S VIEW OF THE WORLD, NOT EVEN AN LF COMPONENT (AS WOULD BE THE CASE IF YOU GOT RID OF THE LF LEVEL AS SUCH, TURNING IT INTO A COMPONENT). SO YOU BASICALLY HAVE TO ASSUME A LESS INTUITIVE NOTION OF CONVERGENCE/CRASHING. I HAVEN'T SEEN THE DETAILS OF THAT NOTION WORKED OUT.

Derivations proceed phase by phase: (18), for example, has the four phases bracketed:

REMEMBER, FOR CHOMSKY THIS ARE JUST STIPULATED PHASES--A DEFINING PROPERTY OF DERIVATIONS BUILT IN IN ORDER TO CUT COMPUTATIONAL COST (NOT EMERGENT, THAT IS, IN THE MSO SENSE). AND KEEP IN MIND THAT THE PHASES IN (18) ARISE SIMPLY BECAUSE YOU TELL THE SYSTEM TO COMPUTE THINGS IN TERMS OF v AND T, HENCE THE FOUR BRACKETS BELOW:

(18) [John [t> thinks [Tom will [t> win the prize]]]]

An alternative that has been suggested{1995 class lectures, and various talks and papers.} is to define phases in terms of convergence.

THIS IS THE DEFINITION I GAVE, ALSO, IN (1998), THE BASQUE PAPER. CHOMSKY IS GOING TO ARGUE AGAINST IT--LET'S SEE THE ARGUMENT.

The two options are then (19):

(19)

(I) Phases are propositional

(II) Phases are convergent

Under (I), LA<{i}> is determined by a single choice of C or v>. Under (II), local determination is not possible.

IT IS CLEARLY TRUE THAT DETERMINATION IS MORE LOCAL IN TERMS OF (I), BUT WHETHER IT IS HOPELESSLY NON-LOCAL (THAT IS, NON-LOCAL IN WAYS THAT HAVE A SERIOUS MATHEMATICAL CONSEQUENCE FOR TRACTABILITY, OR EVEN A NON-REASONABLE SPAN--SAY, HUNDREDS OF SYMBOLS AS OPPOSED TO A COUPLE), THAT DEPENDS ON WHETHER THE SYSTEM HAS A SINGLE LF LEVEL OR RATHER THE SYSTEM ACCESSES INTERPRETATION IN CASCADES, THUS THROUGH A COMPONENT--NOT A LEVEL--OF LF. NEEDLESS TO SAY, WHETHER THE SYSTEM AT LARGE CAN BE MADE TO WORK WITHOUT AN LF LEVEL (BASICALLY, WITHOUT GLOBALITY) IS SOMETHING THAT HAS TO BE ARGUED FOR INDEPENDENTLY. BUT MY POINT IS NOW SIMPLY THIS: THE LOCALITY ISSUE IN (19) DEPENDS ON WHAT THE ANSWER TO THAT QUESTION IS. IF THE SYSTEM ONLY HAS AN LF COMPONENT, THE DIFFERENCES IN LOCALITY IN (19) ARE

NEGLIGIBLE.

Complexity considerations therefore favor Option (I), and again the empirical evidence supports the same conclusion.

WELL, LET'S SEE ABOUT THAT NEXT.

The two options have similar (though not identical) consequences in such cases as (18), but are clearly distinguished elsewhere. One case is A'-movement, as in (20)

{An island effect, if it exists at all, is very weak with such structures.}:

I'M OLD ENOUGH TO HAVE SEEN SENTENCE (20) BE A BARRIERS VIOLATION! BUT ANYWAY, LET'S ASSUME THE SENTENCE IS GOOD TO SEE WHETHER THIS IS A KIND OF EXAMPLE THAT WE SHOULD USE TO DECIDE ON THE PROPER DEFINITION OF PHASE.

(20) which article is there some hope [<"> that John will read t>{wh->}]

For reasons to which we return, assume that the wh>-phrase has an uninterpretable feature analogous to structural Case for nouns, which requires it to move to its final position in an appropriate C.

WELL, I'M SURE YOU RECALL A VERY DIFFERENT STORY IN CHAPTER 4, WHERE IT WAS ARGUED THAT WH-PHRASES ONLY HAVE INTERPRETABLE FEATURES. THE REASONS FOR THIS WERE TWO. ONE WAS CONCEPTUAL: IT WAS TAKEN AS GIVEN THAT THE (PRESUMABLY) D FEATURE OF A WH-ELEMENT IS INTERPRETABLE BECAUSE IT HAS SOME SEMANTIC IMPORT. MORE IMPORTANTLY, THOUGH, THE SECOND REASON WAS EMPIRICAL, AND HAD TO DO WITH WELL KNOWN EXAMPLES OF THE SORT IN (i):

(I) WHO SAW WHAT

IF WH-ELEMENTS HAVE UNINTERPRETABLE FEATURES, THEN WHY IS (I) EVER POSSIBLE? THERE IS, MIND YOU, A WAY OUT--BUT ONE THAT POSES AS MANY QUESTIONS AS IT ANSWERS. IT COULD BE THAT STRONG FEATURES ARE NOT INTRINSIC TO CATEGORIES, AND THAT FOR SOME REASON THEY HAVE TO BE ASSIGNED TO ITEMS IN NUMERATIONS. IF FURTHERMORE YOU ASSUME THAT IN LANGUAGES LIKE ENGLISH YOU CAN ONLY HAVE ONE STRONG WH-FEATURE PER PHASE (OR SOMETHING LIKE THAT, PERHAPS TIED UP TO PRESENCE OF A +WH COMP), THEN YOU MIGHT GET (I) ABOVE. BUT MANY PROBLEMS THEN ARISE. TO BEGIN WITH, IF WE'RE GOING TO DO THINGS THIS WAY, WHAT FORCES A NORMAL DP TO HAVE, SAY, CASE? OR ARE WE GOING TO SAY THAT CASE, UNLIKE WH, *IS* OBLIGATORY TO DP? THAT COMES CLOSE TO RESTATING THE PROBLEMS... AT ANY RATE, LET'S CONTINUE, ASSUMING CONTRA THE CHAPTER 4 TREATMENT THAT WH-PHRASES HAVE STRONG FEATURES.

Then " is a phase under Option (I) but not Option (II) (it does not converge, containing an uninterpretable feature). The only phase is (20) itself; merger of there> blocks raising of John>

to SPEC-TP within ", so (20) is underivable without look-ahead. That is unnecessary under Option (I), with " derived from LA<{i}> lacking EXPL.

IN OTHER WORDS, WE FALL BACK INTO THE MARANTZ/ROMERO PUZZLE. THEREFORE YOU CANNOT DEFINE PHASES AS THE MINIMAL AMOUNT OF STRUCTURE YOU NEED TO CONVERGE, FOR IN THESE INSTANCES THE STRONG WH-FEATURE FORCES YOU TO LOOK UP TO THE NEXT PHASE IN ORDER TO FIND THE NECESSARY MATERIALS FOR CONVERGENCE. BUT HOLD ON: THE CRUCIAL ARGUMENT IS BASED ON A DUBIOUS ASSUMPTION! IN THE CHAPTER 4 TERMS THE ARGUMENT DOESN'T HOLD AT ALL. AT THE SAME TIME, THE REAL QUESTION HERE IS WHAT IS BEHIND SUCCESSIVE CYCLIC WH-MOVEMENT. THE DEFINITION OF PHASE IN TERMS OF CONVERGENCE DEMANDS THAT YOU TAKE A POSITION ON THAT. IT IS CONCEIVABLE THAT WHATEVER TAKES A WH-PHRASE TO AN INTERMEDIATE COMP IS ENOUGH TO DEFINE ITS CONVERGENCE PROPERTIES. THE DEVIL IS IN THE DETAILS THOUGH, AND WE HAVE NO CLUE AS TO WHAT TAKES A WH-ELEMENT TO AN INTERMEDIATE COMP. WE RETURN TO THIS.

The descriptive typology of movement, a leading research topic for years,

{Early work sought to establish the categories of A- and A'-movement ("Move-NP," "Move-wh>-"), later head movement, while parallel inquiries sought commonalities. Important outcomes were Rizzi's theory of relativized minimality and the Lasnik-Saito Move-" theory. The distinctions mentioned here cross-cut these categories.}

offers other reasons to suspect that phases are real, understood under Option (I). There are several categories: movement can be feature-driven or not; and in the former case can be directly or indirectly feature-driven. Typical cases include raising to subject (directly feature-driven), the non-final stages of successive-cyclic movement (indirectly feature-driven), QR and "stylistic movement" (perhaps not feature-driven).

{"Stylistic" operations might fall within the phonological component (see MP>, chap. 4.7.3; Kidwai 1996).

THIS IS SIMILAR TO AN IDEA IN HOFFMAN (1996), FOR WHOM STYLISTIC OPERATIONS INVOLVED ABSENCE OF MERGE IN THE OVERT COMPONENT, AND THEREFORE ABSENCE OF ORDER IN TERMS OF SYNTACTIC TOOLS.

Operations lacking overt counterparts and apparently not interacting with C<HL> might be among the principles of interpretation of LF, hence "post-cyclic,"

THIS IS AN INTERESTING OPENING WHICH I'LL RETURN TO WHEN WE DISCUSS HORNSTEIN AND URIAGEREKA 1999 (HENCEFORTH H&U), WHERE PURE LF PROCESSES (E.G. POLARITY LICENSING) ARE ANALYZED.

inspecting a representational level

AGAIN THE ASSUMPTION HERE (THOUGH PERHAPS NON-TECHNICALLY) IS FOR A LEVEL. IT'S CLEAR THAT YOU NEED SOME SORT OF UNIFIED OBJECT TO DO, SAY,

BINDING CONDITION C. STILL, WHETHER THIS CAN BE ACHIEVED WITH, FOR INSTANCE, A FILE SYSTEM (THAT IS, A SET OF LINKED SUB-COMPONENTS) AS OPPOSED TO A SINGLE LEVEL IS AN ENTIRELY OPEN QUESTION.

in the manner of many other systems (including Binding Theory, on the assumptions of MP>). If so, much of the very enlightening recent work on ellipsis and antecedent-contained deletion (along with event structure and other topics) could be understood as an exploration of the language-external systems at the border of the language faculty, roughly analogous to acoustic and articulatory phonetics at the sound side.}<TYP = footnote>

THAT'S ESSENTIALLY THE LINE I'LL BE PRESENTING WHEN I DISCUSS H&U.

Indirect feature-driven movement (IFM) subdivides into types depending on the attracting head H in the final stage: (I) A-movement when H has N-features (yielding the Case-agreement system), or (II) A'-movement when H has P>-features> of the peripheral system (force, topic, focus...).

REMEMBER, THIS IS *INDIRECT*, THUS NOT THE FINAL STAGES OF THE MOVEMENT, BUT IN ESSENCE SUCCESSIVE-CYCLIC STEPS OF A LARGER MOVEMENT.

{The categories might overlap, but unproblematically it seems. System design should preclude unwanted cases of improper movement.

IT'S ACTUALLY A BIT STRONGER. SYSTEM DESIGN (OR ELSE VERY GOOD OUTPUT CONDITIONS) SHOULD *MOTIVATE* WANTED CASES OF PROPER MOVEMENT AS WELL...

That seems attainable, but must be demonstrated. I will continue to restrict attention to raising of XP (see p. 00).} The intuitive argument for IFM has always been that locality conditions require "short movement" in successive stages, leading to convergence in the final stage.

THAT WAS, YOU MAY RECALL, THE WHOLE IDEA BEHIND THE SUBJACENCY CONDITION, AS SEEN IN LONG-DISTANCE MOVEMENT ACROSS AN 'ESCAPE HATCH'.

We can express a version of this idea as a "phase-impenetrability condition," strengthening further the notion of cyclic derivation. Given HP = [" [H ß]], take ß to be the domain> of H and " (a hierarchy of one or more SPECs) to be its edge>. The thesis under consideration is (21):

> (21) In phase " with head H, the domain of H is not accessible to operations outside ", but only H and its edge

ONCE AGAIN THIS RELATES TO THE DISCUSSION ABOVE CONCERNING THE OPACITY OF PHASES AND WHY IT SHOULD ARISE. HERE THE SUBSTANTIVE CLAIM IS THAT ONLY THE *DOMAIN* (CAREFUL AGAIN WITH TERMINOLOGY, SINCE THIS USED TO BE CALLED THE internal domain IN CHAPTER 4) OF A HEAD IS OPAQUE, WHILE THE EDGE (WHAT USED TO BE THE checking domain) IS TRANSPARENT. THIS IS GOING TO WORK, YES, BUT WE SHOULD REALLY BE

ASKING WHY. NOTE THAT A SIMILAR QUESTION IS POSED FOR S-MOVEMENT, SAY IN THE N&U SYSTEM. BASICALLY, YOU WANT TO ALLOW THE EXTRA STEP (IN THIS CASE OF S-MOVEMENT) PRECISELY TO ALLOW LONG-DISTANCE STUFF. BUT HOW DO YOU PREVENT IT IN UNWANTED INSTANCES? WE RETURN TO THIS EXTENSIVELY.

The cycle is so strict that operations cannot "look into" a phase " below its head H. H itself must be visible for selection and head-movement, hence its SPECs as well.

THE *HENCE* BIT IS SOMEWHAT OPAQUE. IT DEPENDS ON THE BRACKET NOTATION {H, {{H, ...}, SPEC}} WHERE H IS A LABEL AND {H,...} AND SPEC ARE THE TERMS. YOU COULD ARGUE THAT NOTHING ELSE IS ACCESSIBLE TO THE SYSTEM ACROSS PHASES. AS YOU'LL SEE, THIS IS SOMEWHAT MORE NATURAL IN TERMS OF S-MOVEMENT, BUT IT IS UNDERSTANDABLE IN PHASE-TERMS AS WELL (THE INTUITION BEING THAT STUFF INSIDE A PHASE ENTERS INTO THE PARA NOTION WITH REGARDS TO OTHER PHASES, HENCE IS NOT REALLY 'IN-CONSTRUCTION-WITH' STUFF OUTSIDE THE PHASE; THE BRACKET NOTATION CAPTURES THAT TELLING YOU THAT THE SYSTEM ONLY SEES THE TOP OF THE STRUCTURE--LABEL AND TERMS--IN PARA CONDITIONS).

Condition (21) yields a strong form of subjacency.

{It also suggests a new approach to some ECP issues, e.g., subject extraction (Idan Landau, pc).}

INDEED THIS IS TRUE (FOR THAT-T EFFECTS), BUT AGAIN, THE DEVIL IS IN THE DETAILS...

For A-movement, it should follow from the theories of Case-agreement/locality.

{To clarify this and related conclusions and establish them in full generality requires a far more comprehensive review and analysis than is undertaken below. Similar qualifications hold throughout.}<QUAL = footnote>

The stipulation is for clausal A'-movement, the basic question from the earliest study of these topics. We return to some speculation about reducibility to economy conditions. The intended consequences do not follow if phases are determined by convergence,

THIS SEEMS A BIT STRONG, AND DEPENDS ON FACTORS I'VE ALREADY MENTIONED. I WILL NOT REPEAT THIS ANY LONGER AND WILL RETURN TO THE ALTERNATIVE LATER ON IN THE SEMESTER.

but the propositional alternative (Option (I) of (19)) accommodates them, with verbal phases limited to transitive v> with N-features and external argument.

YOU WILL SEE THAT CHOMSKY DOESN'T WANT EXTRA PHASES, FOR INSTANCE, IN A RAISING PASSIVE--WHERE <u>v</u> WOULDN'T BE PRESENT.

The impenetrability condition requires that A'-movement target the edge of every phase, CP and Vp>. There is evidence from reconstruction-effects and parasitic gap constructions that this may be true.

{Fox (1998), Nissenbaum (1998). If adjunction is restricted as suggested in MP> (chap. 4.7.3), then movement to the edge will be to a SPEC position for v>P as well as CP.

THIS IS VERY REMINISCENT OF A SIMILAR IDEA IN THE BARRIERS FRAMEWORK, BUT REMEMBER, EACH OF THOSE INTERMEDIATE STEPS WILL HAVE TO BE JUSTIFIED...

Phases might also be the target for QR, if this noncyclic operation targets C', merging the raised quantifier phrase between C and SPEC-C;

THIS IS A REFERENCE TO RICHARD'S TUCKING-IN, AS IS THE FOLLOWING:

see note 108.}<QURA = footnote>

The idea that IFM applies only if needed to guarantee eventual convergence appears to raise questions of look-ahead. These are obviated if (21) holds. Local determination is straightforward: and an uninterpretable feature in the domain determines at the phase level that the derivation will crash.

YOU NOW SEE WHY CHOMSKY CANNOT MAINTAIN THE CHAPTER 4 VERSION OF WH-MOVEMENT, WITH AN INTERPRETABLE FEATURE IN THE WH-PHRASE... NOTE, IN ANY CASE, THAT SIDESTEPPED HERE IS WHY MOVEMENT TO THE EDGE IS POSSIBLE (IN TERMS OF GREED). OF COURSE, A SIMILAR ISSUE ARISES FOR S-MOVEMENT: WHY IS IT POSSIBLE WHEN NEEDED FOR LONG-DISTANCE CONSIDERATIONS?

{Convergence is not guaranteed, of course (it can fail in many ways); only permitted without look-ahead, the desideratum we are exploring.

AND RECALL, ALSO, THAT THE NOTION OF CONVERGENCE AS THE DERIVATION GOES BY (WHICH AS WE SAW IS NON-TRIVIAL) IS CRUCIAL HERE.

Conditions could be added to restrict crash, but they are redundant, simply restating properties of convergence, unless motivated in some other way. Questions arise about operations that appear to violate subjacency (see references of note 93, among others).}

Let us return to the basic configuration (4) for CFCs, repeated as (22), with XP the extra SPEC determined by the EPP-feature of the attracting head H:

(22) " = [XP [(EA) H YP]]

examples of (22) are raising to subject (yielding (23A)),

(23)(A) XP - [T YP]

Object Shift (OS, yielding (B), with XP = DO and t> its trace),

(B) XP - [SU [v> [V t>]]]

and overt A'-movement (yielding (C), with H = C and XP a wh>-phrase (see note <LIGHT>)):

(C) XP - [C YP]

The EPP-feature of T might be universal. For the phase heads v>/C, it varies parametrically among languages and if available is optional. Cases (B) and (C) are similar, and unlike (A), in other respects. One is that the EPP-feature can be satisfied by Merge of an expletive EXPL in (A), but not in (B)/(C) (see (5i)).

ALTHOUGH, AGAIN, whether, ETC...

The fact might be unexpected for (B), because (B) and (A) enter into the Case-agreement system in much the same way; and the relation of SPEC-H, H, and a related phrase ß in the complement of H appears to be similar for T and v>. Phases are determined by a choice of C/v>, not T, which suggests a basis for the similarities and asymmetry. The fact that the EPP-feature when available is optional for C/v> suggests that it is a property of the phase PH:

NOTE, INCIDENTALLY, THAT MY SKEPTICISM ABOVE WAS NOT ABOUT HAVING PHASES AS C/v, BUT RATHER HAS HAVING THAT FOLLOW FROM SOMETHING AS UNSPECIFIED AS 'PROPOSITIONAL STUFF'. THAT MEANS I AM PERFECTLY HAPPY WITH (24) BELOW, AND I'D LIKE TO UNDERSTAND *WHY* PHASES HAPPEN TO BE WHAT THEY ARE. AT ANY RATE, THE IDEA THAT PHASES ARE DOMAINS FOR P AND EPP FEATURES (REMEMBER P MEANS 'PERIPHERAL') SEEMS TO ME TO BE VERY REASONABLE, BUT AGAIN GETS US CLOSER TO A D-STRUCTURE COMPONENT... AFTER ALL, NOW PHASES ARE DOMAINS THAT CORRESPOND TO THE OLD THETA-DOMAINS, ETC.

(24) The head H of phase PH may be assigned an EPP- and P-feature

{Parametrically varying properties of H enter into the application of (24), which might be extended to head-movement (see p. 00 and note 93). The P-feature should be redundant, a reflex of the EPP-feature if H does not already have an appropriate P-feature (say, the Q-feature of interrogative C).

THIS IS A CRYPTIC NOTE THAT I DO NOT FULLY UNDERSTAND. I SUSPECT THIS IS ALL TO 'MOTIVATE' THE NECESSARY, YET UTTERLY MYSTERIOUS INTERMEDIATE STEPS OF SUCCESSIVE CYCLIC MOVEMENT, WHICH SEEM TO BE THERE JUST SO THAT THE SYSTEM SATISFIES LOCALITY, WITH NO PARTICULAR CONCERN FOR CONSIDERATIONS ABOUT GREED. MIND YOU, THE IDEA THAT A P-FEATURE SHOULD DRIVE INTERMEDIATE MOVEMENT IS INTERESTING, AND SUGGESTED, FOR INSTANCE, IN GROHMAN'S AND WU'S RECENT THESES--BUT IT IS HARD TO SEE THAT, SAY, TOPICALIZATION IS WHAT GEARS INTERMEDIATE MOVEMENT IN GENERAL, AS WE WELL SEE LATER ON IN THE SEMESTER.

The two features are introduced to allow the general theory of movement to apply without change in this case.}

THAT, OF COURSE, IS WORRISOME, SINCE THE FEATURES APPEAR TO THEN BE MERE NOTATION TO GO ON WITH THE SHOW... INDEED, YOU'LL SEE IMMEDIATELY BELOW THAT THIS IS PRECISELY WHAT'S GOING ON.

Once PH is completed, exhausting the lexical subarray from which it is derived, (24) may optionally apply, assigning an EPP-feature to H.

AS FAR AS I CAN SEE THIS JUST MEANS THE EPP-FEATURE IS ENTIRELY OUT OF THE SYSTEM AS SUCH, AND IS PROPOSED IN ORDER NOT TO MAKE MOVEMENT TOO MESSY...

From the strong cyclicity condition that renders H inert beyond the phase itself (see (17)), it follows that EPP must be satisfied by raising within PH: pure Merge from outside PH is barred. In (23B,C) XP is raised from within the domain of H = v>/C, completing the account for case (i) of (5).

NOTE THAT, IN EFFECT, WHAT PREDICTS EXPLETIVES ONLY IN T IS THE FACT THAT T IS NOT A PHASE, HENCE YOU COULD HAVE MERGED IN THE CRUCIAL POSITION OF THE EXPLETIVE. IN THE SPEC OF v OR C THE MERGE OPTION DOESN'T EXIST BECAUSE YOU'VE EXHAUSTED THE ELEMENTS IN THE CYCLIC ACCESS TO THE NUMERATION AND BY FIAT (24) APPLIES *AFTER* THAT, INTRODUCING THE EPP FEATURE AT THAT POINT--SO YOU'RE ONLY ALLOWED TO MOVE TO SATISFY THAT FEATURE.

{In MP>, an unsatisfactory argument was required to bar expletives from merging in OS constructions. The analogous problem arose for CP, but was ignored.}

ALTHOUGH REMEMBER WHAT I KEEP REPEATING, whether DOES HAVE SOME OF THE PROPERTIES OF AN EXPLETIVE: IT IS MERGED AND IT ASSOCIATES TO THE CLAUSE AS IF IT WERE ITS 'ASSOCIATE'...

The remaining properties of (5) should follow from the theories of Case-agreement and locality to which we turn shortly. The picture might be extended to incorporate QR, if alongside the N- and P-features that drive movement, there are also QU-features, attracting quantifiers that pied-pipe an appropriate phrase.

THAT'S THE BEGHELLI LINE, AND THE FOLLOWING ONE, THE REINHART/FOX IDEA.

One might also explore a generalization of the idea that operations can apply only if they have an effect on outcome; see p. 00. This would translate here into restricting (24) to the condition in which it permits IFM or specific interpretations associated with peripheral positions: e.g., specificity and informational conditions on OS).

{For some speculations along similar lines, see MP>, condition (76) of chap. 4.5.3, and p. 377.}

There are many problems and ramifications, which I will leave to the side. (24) yields A- or A'-movement depending on whether the phase head has N- or P-features. It might have both. Suppose in the construction (25), all four phase heads are assigned an extra SPEC by (24), associated with P-features for C and v><2> but not for v><1>:

(25) SPEC-C<2>...SPEC-v><2>...SPEC-C<1>...SPEC-v><1>...XP

NOTE THAT THE ASSUMPTION HERE REQUIRES LOTS OF JUSTIFICATION,

INCLUDING THE IDEA THAT THE INTERMEDIATE, HIGHER v HAS SOME P-FEATURES ASSOCIATED TO IT (AGAIN, THAT'S THE GROHMAN/WU LINE).

XP raises through the SPECs in succession, landing finally in SPEC-C<2>. The result is the 4-membered A'-chain (SPEC-C<2>, SPEC-v><2>, SPEC-C<1>, SPEC-v><1>) and the 2-membered A-chain (SPEC-v><1>, XP) (formed by Object Shift). SPEC-v><2> is an A'-position, by virtue of the P-feature associated with the extra SPEC introduced by (24); v><2> also had N-features involved in object Case/agreement but these would have been deleted phase-internally before (24) assigns the extra SPEC.

{On some assumptions, IFM passing through SPEC-v><2> is improper movement, though not here.}

There are more complex cases to consider, along with a variety of other issues.

{Among other questions, what is the status of small clauses, or relative clauses and other adjuncts? Possibly the latter are derived "in parallel," in the manner of multidimensional analyses of coordination or parentheticals, with their own Las and the ultimate status of the adjunct determined in the larger structure in which it is inserted (as for other multidimensional structures).}

THE MULTIDIMENSIONAL ANALYSIS OF CONJUNCTIONS IS PURSUED BY GOODALL 1984. THIS VIEW OF ADJUNCTS WOULD HAVE THE ADVANTAGE OF ELIMINATING LINEARIZATION/SCOPE PROBLEMS FOR THEM, OF THE SORT NOTED BY ERNST AND OTHERS, BUT I WILL RETURN TO THESE ISSUES LATER IN THE SEMESTER.

But it seems that the cyclic approach to accessing lexical arrays is plausible on conceptual and empirical grounds, along with Option (I) of (19) and the condition (24) on extra SPECs.

HERE STARTS A NEW TOPIC, CENTRAL TO VARIOUS OTHERS (E.G. S-MOVEMENT OR THE DERIVATION OF THE LCA IN MSO).

Nothing has been said yet about the fact that C<HL> forms syntactic objects in parallel, according to the derivational approach adapted from MP>. Consider, for example, the expression (26):

> (26) the demonstration that glaciers are receding showed that global warming must be taken seriously

The pre-final phases of the derivation are the syntactic objects corresponding to (27)

{See note <LIGHT>. Many questions about the internal structure of the words are put aside.}:

> (27)
>
> (i) P<1> = [<CP> that global warming must be taken seriously]
>
> (ii) P<2> = [<CP> that glaciers are receding]

(iii) P<3> = [<v>P> [the demonstration P<2> [show P<1>]]]

NOTICE, INCIDENTALLY, HOW CLOSE (WITH APPROPRIATE ABSTRACTIONS) THESE UNITS ARE TO KERNEL SENTENCES IN THE PRE ASPECTS SYSTEM...

For each new phase, a subarray provides the lexical material required and the operations proceed in the manner already sketched, with P<1>/P<2> unordered. Step (iii), for example, is formed from the subarray {the>, demonstration>, show>}; repeated Merge yields DP = [the> [demonstration> P<2>]] and then P<3> = [DP [show> P<1>]], with theta roles assigned. The next subarray adds T and C, permitting the derivation to continue to P<4> = (26).

IN THE EXAMPLE CHOMSKY CHOSE IT IS PLAUSIBLE TO SPEAK NOT JUST OF PARALLEL MERGER, BUT ALSO OF PHASES, AS WE SAW. HOWEVER, PARALLEL MERGER ALSO ARISES IN THE SIMPLE 'THE MAN SAW A WOMAN', ALTHOUGH THERE IT IS LESS OBVIOUS THAT YOU ALSO HAVE DIFFERENT PHASES (FOR SURE YOU DON'T HAVE EXTRA C'S OR v'S, ALTHOUGH WE DON'T KNOW WHAT CHOMSKY THINKS OF D AS A POSSIBLE PHASE--IN FACT THE ISSUE IS EXTREMELY TRICKY).

We have found evidence that computational complexity enters into language design, but one might ask whether the argument can be strengthened on purely conceptual grounds. Consider theses (I)-(IV), where (II-IV) presuppose a positive answer to (I):

(28)

(I) Computational complexity matters for a cognitive system

(II) The solution must be comprehensive, with a guarantee of

"quick decision" for all derivations

(III) Complexity should not be allowed to grow "too fast"

(IV) Decisions in computation attend only to principles of UG

In recent discussion, such theses are sometimes adopted as virtual conceptual necessity.

{See, e.g., Collins (1997), Johnson and Lappin (1997), Yang (1997).}<ECON = footnote>

That is a questionable move, however.

QUICK COMMENT HERE. WHAT IS A DUBIOUS MOVE IS THAT COMPUTATIONAL COMPLEXITY ITSELF, PARTICULARLY IN THE TERMS IN (28) (WHICH DO NOT CORRESPOND TO ANY INDEPENDENT THEORY OF COMPLEXITY) SHOULD BE CONCEPTUALLY NECESSARY. QUITE A DIFFERENT THING, HOWEVER, IS THAT STRUCTURAL ECONOMY SHOULD BE CONCEPTUALLY NECESSARY. THAT, IN OUR UNIVERSE, SEEMS IF NOT NECESSARY AT LEAST THE ONLY GAME IN TOWN (D'ARCY THOMPSON'S THESIS).

One reason is that the theses seem to presuppose that the derivational approach is correct, which

is hardly obvious (if the question is even meaningful).

THE QUESTION IS PROBABLY MEANINGFUL, BUT VERY HARD TO ESTABLISH. NONETHELESS, IF THE DERIVATIONAL APPROACH *IS* CORRECT YOU DO EXPECT CONSIDERATIONS ROUGHLY OF THIS SORT.

That aside, it is not clear why (I) should be true. Thesis (II) raises further questions. Languages L have expressions that satisfy all conditions of UG and L but are "unusable" for some reason; see note <USE>. It could turn out that among these are expressions that cannot be derived efficiently, a result that would be interesting if true (see Chomsky 1991). Interpretation of Thesis (III) in the technical sense of complexity theory also calls for justification. Why should we expect such properties to be relevant to natural language?

THAT'S A FAIR POINT, AND HAVING THOSE EXPECTATIONS MAKE SENSE ONLY IF YOU HAVE FURTHER EXPECTATIONS ABOUT LANGUAGE BEING COMPUTATIONAL, OR SOME SUCH THING (IN FACT A CLASSICAL COMPUTER, NOT A QUANTUM ONE, WHERE THESE CONSIDERATIONS ARE PRESUMABLY BESIDE THE POINT).

Thesis (IV) is intended to bar look-ahead by requiring that at each choice point in a derivation, an irrevocable decision must be made in terms of principles of UG. The intuitive idea is that only "fundamental principles" of UG can be invoked, not consequences of these principles (however easily determined). Again, that is not an obvious conclusion.

{The reference to "fundamental principles" in some (obscure) sense is crucial. Otherwise, for properly selected categories of expressions (which may well exhaust the possibilities), look-ahead properties of computation at stage " might be overcome by resort to whatever aspects of UG determine that the wrong choice eventually crashes.}

THIS IS INTERESTING, HARD TO CHARACTERIZE, AND CURIOUSLY RELATED IN AN INDIRECT WAY WITH WHAT PRESUMABLY GOES ON IN LANGUAGE ACQUISITION, WHERE THE CHILD IS NOT ALLOWED TO MAKE EXTREMELY INDIRECTLY JUDGEMENTS ABOUT STRUCTURE (SAY IN A CUE-BASED SYSTEM OF THE SORT LIGHTFOOT HAS BEEN ARGUING FOR), AND MUST INSTEAD RESTRICT ALL DECISIONS TO VERY LOCAL AND VERY BASIC UG STUFF. IT'S AS IF ALL UG GAMES WERE PLAYED LOCALLY (ALTHOUGH QUESTIONS STILL REMAIN, OF COURSE, ABOUT HOW LOCAL IS LOCAL--WHAT ARE THE DOMAINS IN ALL THIS, WHICH NATURALLY RELATES ALSO TO DEGREE-0 IN LEARNABILITY).

Theses (III) and (IV) are often held to undermine the account of (9)-(10)/(12)/(15)/(16) reviewed earlier on the grounds that it involves look-ahead.

I NEVER UNDERSTOOD THAT CHALLENGE, SINCE THE LOOK-AHEAD IN THOSE INSTANCES SEEMS COMPLETELY TRIVIAL, PARTICULARLY WHEN WE SEE THAT 'GLOBAL' CONSIDERATION REDUCE TO CYCLES...

If that were true, the validity of the theses would be put into question, not the analysis, unless

some justification can be found for them.

YEAH, BUT I DON'T EVEN SEE THAT THE THESES ARE QUESTIONED BY THE CHALLENGE.

Notice how difficult it is even to give a clear formulation of thesis (IV), hence of the argument that the proposed analysis of these constructions even raises look-ahead issues.

{The issue is whether crash "frees up" derivational paths not selected earlier, as assumed in MP> (incorrectly, we assume here).

THIS IS A MAJOR CHANGE, FOR BETTER OR FOR WORSE. ALTHOUGH IT IS TRUE THAT THE SYSTEM THAT ALLOWS SUCH A "FREEING UP" IS COMPUTATIONALLY MORE COMPLEX--OBVIOUSLY--IT IS REALLY NOT OBVIOUS AT ALL THAT WE SHOULD GET RID OF IT JUST FOR THAT REASON. AGAIN, FIRST OF ALL IT IS NOT OBVIOUS THAT COMPLEXITY HERE IS "UNFATHOMABLE", OR IN FACT SIGNIFICANT IF THE SYSTEM IS CHUNKED DOWN THE WAY WE ARE SUGGESTING, WITH CYCLES AND ALL THAT. AND SECOND, A SYSTEM THAT EVALUATES A CLASS OF ALTERNATIVES MIGHT BE NECESSARY ON DEEPER GROUNDS--FOR INSTANCE, TO DEAL WITH THE PHENOMENON OF PREFERENCE, WHICH I RETURN TO. GRANTED, IF YOU CAN GET AWAY WITHOUT THE MORE COMPLEX SYSTEM, SO BE IT; BUT I DON'T REALLY SEE THE CONCEPTUAL PRESSURE TO GO IN THAT DIRECTION WITHOUT A VERY CAREFUL EXAMINATION OF WHAT'S REALLY COMPLEX AND WHY.

Computations can crash in endless ways, raising no complexity issue. See note 50.}

The assumption is that (IV) is violated by resort to the theta-theoretic principle (6), though apparently not by preference of Merge over Move.

I'M NOT SURE I UNDERSTAND THIS SENTENCE. I SUPPOSE IT MEANS "THE ASSUMPTION IN THE LITERATURE IS..." I DON'T REALLY UNDERSTAND, EITHER, THE QUESTION BEHIND THE ENTIRE PARAGRAPH. I WOULD HAVE THOUGHT THAT THE CLAIM IS BACKWARDS: THAT IS, (IV) (NAMELY, THAT DECISIONS IN COMPUTATION ATTEND ONLY TO PRINCIPLES OF UG) IS *NOT* VIOLATED BY RESORT TO THE THETA-THEORETIC PRINCIPLE (6), THOUGH APPARENTLY *IT IS* VIOLATED BY PREFERENCE OF MERGE OVER MOVE. AFTER ALL, (6) IS A PRINCIPLE OF UG, I TAKE IT. FURTHERMORE, GIVEN THE FOLLOWING DISCLAIMER, I WONDER WHETHER THERE IS A 'NOT' MISSING IN THE PREVIOUS SENTENCE:

Why is unclear: selection of Merge over Move or conversely is determined (trivially) by principles of UG.

I ASSUME THIS MEANS WHAT I JUST SAID: MERGE-OVER-MOVE IS NOT ITSELF A PRINCIPLE OF UG, ALTHOUGH IT IS DETERMINED BY SUCH PRINCIPLES. AND CONVERSELY:

Furthermore, how do we know that (6) is "a theorem," not a principle of language design (part of

Theta Theory), so that resort to it in fact satisfies the intuition that lies behind this application of (IV)?

HERE CHOMSKY SEEMS TO BE CASTING SOME (REASONABLE) DOUBT OVER THE INTUITION THAT (6) IS, IN FACT, ITSELF A PRINCIPLE. BUT THEN I DON'T REALLY SEE HOW TO INTERPRET THE CENTRAL CLAIM IMMEDIATELY ABOVE WITHOUT THE "NOT" I SUGGESTED. REMEMBER, THIS IS A MANUSCRIPT AND THERE MAY BE SECTIONS WITH TYPOS AND SO ON. AS A MATTER OF FACT, SOME ARE CLEAR IMMEDIATELY BELOW. ANYWAY, THE WHOLE THING ISN'T TOO IMPORTANT, SINCE WHAT'S REALLY IMPORTANT IS THIS:

Again, it seems that we should seek a resolution of the issues on empirical grounds.

My suspicion is that thesis (I) might be correct and perhaps (II), along with versions of (III)-(IV) that restrict choices in derivations to elementary principles of UG and bar even narrowly bounded look-ahead.

THAT'S, OF COURSE, THE BITE OF THIS WHOLE PAPER, ALTHOUGH I SHOULD INSIST THAT NOTHING IN THAT *SUSPICION* PREVENTS US FROM USING MORE COMPLEX COMPUTATIONAL EVALUATIONS, AS I CLAIM ARE NEEDED FOR UNDERSTANDING THE PHENOMENON OF PREFERENCE. THAT'S JUST A FACT (THAT WE CAN SYSTEMATICALLY DETERMINE PREFERENCES).

But the questions are empirical. If theses [SIC] of (28) hold in some form, that would be a surprising empirical fact about language design; evidence is required to establish it. The best evidence I know is two-fold: observed cases seem to support thesis (I) and to be consistent with (II). Inquiry seems to reveal further that postulated properties of language that induce computational complexity were incorrectly formulated, and that when improved, undue complexity is overcome and computation is "local," suggesting that (III) may be valid.

{See Chomsky (1998), Collins (1997), Frampton and Guttman (1998).}

ALTHOUGH I'M NOT GOING TO GO INTO THESE OTHER PIECES IN THE SEMINAR, THESE ARE ALL WORTH READING, AND THEY HAVE CLEARLY INSPIRED MUCH OF THE PRESENT RESEARCH.

Furthermore, there seems no need [SIC] to invoke principles beyond the most elementary ones at any point. If such results are consistently found, they would provide confirming evidence for versions of (28), and for the derivational approach, which provides the framework within which they are formulated.

THAT, I BELIEVE, IS WHAT INTERESTS CHOMSKY THE MOST HERE. REMEMBER, THE COMPUTATIONAL THEORY OF MIND, CLASSICALLY, INVOLVES TWO ASPECTS. ONE IS A DERIVATIONAL ARCHITECTURE (THAT IS, WITH A CERTAIN CHARACTERISTIC SYNTAGMATIC LOGIC) AND THE OTHER A MORE OR LESS DUBIOUS REPRESENTATIONAL BASE (WHERE SYMBOLS AND A PARADIGMATIC LOGIC ARE CODED). I SAY THE LATTER IS DUBIOUS FOR REASONS I MENTIONED, AND ALSO BECAUSE IT IS VERY HARD TO PIN DOWN WHAT THAT IS ALL ABOUT. FOR INSTANCE, ARE THE ELEMENTS THAT ENTER INTO THE CALCULATION OF A

PROTEIN-FOLDING ALGORITHM REPRESENTATIONAL? MOST PEOPLE WOULD SAY NO, BUT IS THE MODEL, THEN, IN SOME SENSE *LESS* COMPUTATIONAL? IT'S NOT TOTALLY OBVIOUS WHAT THAT MEANS, BUT IT WOULD BE NICE TO FIND ARGUMENTS TO THE EFFECT THAT THE SYSTEM IS INDEED COMPUTATIONAL, REGARDLESS OF THE REPRESENTATIONAL ISSUES. ENTERS COMPLEXITY. IF YOU CAN GIVE AN ACCOUNT OF CERTAIN ASPECTS OF THE MODEL THAT ARE EMPIRICALLY NECESSARY AND THAT ARE BASED ON (SERIOUS) COMPLEXITY CONSIDERATIONS, AND FURTHERMORE THOSE CONSIDERATIONS ARE ONLY EXPRESSIBLE IN DERIVATIONAL TERMS, THEN YOU HAVE A CLEAR ARGUMENT THAT THE SYSTEM IS STILL COMPUTATIONAL, REGARDLESS OF THE REPRESENTATIONAL PROPERTIES IT MAY OR MAY NOT HAVE.