

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 [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 SUBSYSTEM 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, TRIVIALY.

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. }

The issues 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 $EXP = \langle PHON, SEM \rangle$, 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 TRIVIAALLY 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 $EXP = \langle PF, LF \rangle$. 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 $EXP\langle 1 \rangle = \langle PF\langle 1 \rangle, LF' \rangle$ and wrong for $EXP\langle 2 \rangle = \langle PF\langle 2 \rangle, LF' \rangle$,

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 $EXP\langle 2 \rangle$, a crucial property distinguishing it from $EXP\langle 1 \rangle$.

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 $EXP = \langle PF, LF \rangle$. 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 $25 = 5 \times 5$ or $5 = \frac{1}{5} \times 25$. 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-specificity [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 intertranslatable, 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 ChI MANIPULATES PROPERTIES, INDEPENDENT OF THE ENTITIES THEY ARE PROPERTIES OF? IF NOT, WHAT IS THE MINIMAL UNIT THAT ChI 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 WOLFIAN (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 THAN 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 (" ,β) and forms K(" ,β) 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 UNDERSTOOD 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 ARCAINE 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 \bar{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 [$\langle \beta \rangle T \langle \beta \rangle \dots$], β 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, $T \langle \beta \rangle$ 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 $v >$, $T \langle \beta \rangle$ agrees with EA, which may raise to SPEC- $T \langle \beta \rangle$ though XP cannot

THE SITUATION HERE IS THE CORE EPP, RELATING T TO THE NEXT \bar{y} 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 \bar{y} 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 \bar{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< β > 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 T_{def}) 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< β > have to do with Case/agreement and EPP. In (iib), if EA does not raise, SPEC-T< β > 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, β in (5ii) should be taken to be the minimal β containing " headed by any CFC H< β >, which would therefore be either T or ν .

IN OTHER WORDS, WE ARE EXPLORING THE MINIMAL β THAT CONTAINS " = [XP [(EA) H YP]], SUCH THAT β 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 " = ν ':

A. ν ' B. C' C. T'

$\wedge \wedge \wedge$

ν CP CP TP T ν P

NORMAL EMBEDDING NORMAL DECLARATIVES NORMAL TRANSITIVES

D. ν P

\wedge

ν TP

RAISING CONTEXTS

WHY CHOMSKY SAYS THAT THEREFORE THE HEAD OF β MUST BE EITHER T OR ν

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< β > 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 TRIVIALY, 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< β > 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 \bar{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 [\bar{c}] what C [\bar{c} TP] John read t]

(ii) there is a possibility [\bar{c}] that proofs will be discovered]

In (7), " = CP with H = C, and the relevant properties of matrix T (T< β >) 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< β > [\bar{c}] [\bar{c} DO] the book] [[\bar{c} SU] many students] [read t<DO>]]

In (8), " is an OS construction with H = \bar{v} and DO = XP of (4)/(5iib). Under Holmberg's generalization V raises to T< β >, which agrees with EA = SU. EA can then raise to SPEC-T< β > ("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< β >-is likely [<"> there to be a proof discovered]

(ii) T< β > [{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< β > 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< β >) depend on " in (9i) but not in (9ii), where it is the head of v>P (not T< β >) that is related to ". The relation of T< β > 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 \bar{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 $LA\langle\{i\}\rangle$ 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 THE 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 $LA_{\langle i \rangle}$ that can be selected for active memory. $LA_{\langle i \rangle}$ 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 \bar{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 SIDWARD (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 \bar{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), $LA\langle\{i\}\rangle$ 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 [$\langle \rangle$ that John will read $t_{\langle \text{wh} \rangle}$]

For reasons to which we return, assume that the $\text{wh} \rangle$ -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 $\text{there} \rangle$ blocks raising of $\text{John} \rangle$

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 $HP = [^{\alpha} [H \beta]]$, take β to be the domain> of H and " α " (a hierarchy of one or more SPECs) to be its edge>. The thesis under consideration is (21):

(21) In phase " α " with head H, the domain of H is not accessible to operations outside " α ", 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 \bar{v} 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 SIDESTEPED 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 β 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 \bar{y} 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 \bar{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) $P\langle 3 \rangle = [\langle v \rangle P] [\text{the demonstration } P\langle 2 \rangle [\text{show } P\langle 1 \rangle]]]$

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 $P\langle 1 \rangle/P\langle 2 \rangle$ unordered. Step (iii), for example, is formed from the subarray {the>, demonstration>, show>}; repeated Merge yields $DP = [\text{the}> [\text{demonstration}> P\langle 2 \rangle]]$ and then $P\langle 3 \rangle = [DP [\text{show}> P\langle 1 \rangle]]$, with theta roles assigned. The next subarray adds T and C, permitting the derivation to continue to $P\langle 4 \rangle = (26)$.

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 \bar{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.

CLARIFICATION

THIS IS A CONTINUATION OF CHOMSKY'S RECENT ARTICLE, WHICH WILL APPEAR IN MARTIN ET AL. (LASNIK'S FESTSCRIPT). AS I SAID BEFORE, THE PAPER IS GLOSSED WITH MY OWN THOUGHTS (IN CAPITALS), AND IT KEEPS FOOTNOTES AS THEY APPEAR IN THE ELECTRONIC MANUSCRIPT. FINALLY, I MUST INSIST ON THE FACT THAT, EVEN IF WE HAVE CHOMSKY'S PERMISSION TO DISCUSS HIS WORK IN THIS FORMAT, THIS IS NOT INTENDED FOR DISTRIBUTION AS SUCH.

<4. The Strong Thesis>

Let us now consider the strongest minimalist thesis (2) and ask where it fails.

PERHAPS IT SHOULD BE EMPHASIZED THAT ONE COULD BE DOING MINIMALISM WITHOUT TAKING THIS STRONG THESIS. FOR EXAMPLE, THE METHODOLOGICAL THESIS THAT HORNSTEIN HAS FORCEFULLY ADVOCATED IS GENUINELY MINIMALISTIC, YET NOT (AT LEAST NOT NECESSARILY) IN THE STRONGEST SENSE THAT WE ARE NOW GOING TO EXPLORE.

The research strategy is to seek "imperfections" of language, properties that language should not have, assuming (2). Apparent imperfections come in several varieties. Given some apparent property P of language, we may find:

(29)

- (i) P is real, and an imperfection
- (ii) P is not real, contrary to what had been supposed
- (iii) P is real, but not an imperfection; it is part of a best way to meet design specifications

The most likely -- hence least interesting -- outcome is (i). Conclusion (ii) is less likely, therefore more interesting; efforts within the minimalist program have sought to establish (ii) in particular cases.

WE WILL REVISIT SOME.

Conclusion (iii) is the most interesting possibility.

WHY THIS IS MORE INTERESTING THAN (II) IS NOT COMPLETELY CLEAR TO ME-- THEY BOTH SEEM EQUALLY INTERESTING!

The question is novel, perhaps premature; or unreasonable, because the strongest minimalist thesis (2), or even weaker versions, is not correct, as one might expect for some complex biological system.

AGAIN, THIS IS THE IMPORTANT THEME ABOUT BIOLOGICAL MESSINESS THAT WE SAW AT THE BEGINNING.

Adopting (2) as a point of departure, assume that FL provides no machinery beyond what is needed to satisfy minimal requirements of legibility and that it functions in as simple a way as possible. We would like to establish such conclusions as (A)-(D):

{These are conceptual "good design" conditions, but fairly trivial ones, based on the assumption that less machinery is better than more.}

(30)

(A) The only linguistically significant levels are the interface levels

OF COURSE, THE PROBLEM HERE IS TO KNOW WHAT ARE REALLY THE INTERFACE POINTS, SOMETHING WHICH IS NOT USUALLY DISCUSSED. (SOUND AND MEANING IS *OBVIOUSLY* TOO SIMPLISTIC.)

(B) The <interpretability> condition: LIs have no features other than those interpreted at the interface, properties of sound and meaning

SAME ISSUE--WHAT ARE REALLY THE INTERFACES?

(C) The <inclusiveness> condition: No new features are introduced by C<HL>

THIS IS WHAT I CALLED BEFORE THE CONSERVATION LAW. OF COURSE, THE CONSERVATION LAW INCLUDES *BOTH* THE INCLUSIVENESS CONDITION AND A CONDITION PREVENTING UNRECOVERABLE DELETIONS. NUNES (1995) CLAIMS THAT THE CLASSICAL RECOVERABILITY CONDITION FOLLOWS FROM ECONOMY, UNDER THE ASSUMPTION THAT RECOVERABILITY RULES ONLY OVER DELETION, AND IF YOU CAN YOU SHOULD TRY TO AVOID DELETION (IN OTHER WORDS, IF YOU DO NOT DELETE, YOU DO *LESS* OPERATIONS THAN IF YOU DO DELETE). THAT SEEMS TO ME TO MISS THE POINT. IF I WERE TO BE ALLOWED TO DELETE THE ENTIRE SET OF WORDS IN THE NUMERATION (SAY, AFTER SELECTING THEM AND EITHER BEFORE OR AFTER MERGING THEM) I'D OBVIOUSLY SIMPLIFY THE SYSTEM: I'D DO NOTHING BEYOND THAT POINT, NO MOVE, NO SPELL-OUT, NO ATTRACT, NOTHING. DELETION WOULD BE EXTREMELY OPTIMAL--IT WOULD REDUCE THE DERIVATION TO SELECT, PERIOD. OBVIOUSLY THAT'S ABSURD, AND WHAT PREVENTS IT IS THE TACIT ASSUMPTION THAT IF John IS IN THE NUMERATION IT SHOULDN'T BE DELETED, IT HAS TO MAKE IT TO LF. THAT'S RECOVERABILITY. WHY THE SYSTEM SHOULD HAVE RECOVERABILITY, SINCE AS WE SAW DOESN'T FOLLOW FROM ECONOMY, IS THEN INTERESTING--AS IS WHY THE SYSTEM SHOULD HAVE INCLUSIVENESS. MY SUGGESTION IS THAT WE THINK OF THOSE TWO TOGETHER AS A CONSERVATION LAW. CONSERVATION LAWS ARE THE MOST FUNDAMENTAL IN THE NATURAL WORLD. IN PHYSICS THEY STARTED WITH THE LAW OF CONSERVATION OF MATTER, THEN ENERGY, THEN MATTER-ENERGY, AND THEY NOW EXTEND TO SOME SEVEN OR MORE CONSERVATION LAWS IN ALL REALMS OF REALITY (E.G. CONSERVATION OF A PARTICLE SPIN, AT THE BASIS OF FIELDS, AND SO ON). EVEN IN OTHER FIELDS

YOU COULD SPEAK OF CONSERVATIONS; FOR INSTANCE, THE INSTINCT TO SURVIVE, WHATEVER THAT IS, IS THE ENGINE THAT DRIVES EVOLUTION-- WITHOUT IT, THE WHOLE LOGIC OF THE SYSTEM COLLAPSES. THE FACT THAT THESE CONSERVATIONS ARE CALLED LAWS CLEARLY TELLS YOU THAT NOBODY REALLY KNOWS WHERE THEY COME FROM, OR HOW TO REDUCE THEM TO SOMETHING MORE BASIC. IN ANY CASE, BEHIND THE CONSERVATION OF INFORMATION THAT I'M TALKING ABOUT HERE IS THE IDEA THAT PATTERNS OF FORM ONLY EMERGE IN THE DYNAMIC INTERACTIONS OF MORE ELEMENTARY FORMS. TAKE FOR INSTANCE AN EDDIE THAT FORMS IN A CURRENT: WE KNOW IT TO BE THE RESULT OF THE MACROSCOPIC INTERACTION OF FLUID MOLECULES THAT FEED-BACK INTO THE SYSTEM AFTER SOME OBSTACLE. NOTHING IN THE SYSTEM IS *ADDED* SO THAT YOU CAN GET THE EDDIES FORM; THEY ARISE OUT OF THEIR OWN BUMPS AND MOVES, NOTHING ELSE; LIKEWISE, YOU CANNOT CHOSE NOT TO HAVE THE EDDIES IF THE RIGHT OBSTACLES AND FLUID VISCOSITY AND VELOCITY ARE IN PLACE--THAT INFORMATION SIMPLY DOESN'T GO AWAY. YOU CAN SPEAK OF CONSERVATION PROPERTIES HERE BECAUSE THE RESULTING PATTERN (WHAT COMES OUT) IS A CONSEQUENCE OF WHAT CAME IN. I PROPOSE THAT WE EXTEND THE ANALOGY TO THE INCLUSIVENESS/RECOVERABILITY PROPERTIES OF DERIVATIONS, WHICH WOULD THEN FIT NICELY WITH OTHER SYSTEMS WHERE FORM EMERGES OUT THERE. AND INCIDENTALLY, IF YOU'RE TEMPTED TO THINK THAT PHRASES OR CHAINS ARE CREATED OUT OF NOTHING IN THE SYSTEM, THINK AGAIN: THEY'RE CREATED ALRIGHT, BUT NOT OUT OF NOTHING; THEY ARE PRECISELY LIKE THE EDDIES: CREATED OUT OF MACROSCOPIC INTERACTIONS OF PARTICLES. CHAINS, SAY, ARE CREATED OUT OF INTERACTIONS OF PHRASES, WHICH ARE CREATED OUT OF INTERACTIONS OF WORDS. THE KEY IS NOT TO ADD ANYTHING TO THAT--THAT'S A CONSERVATIVE SYSTEM, VERY MUCH IN LINE WITH THE STRONG MINIMALIST THESIS.

(D) Relations that enter into C<HL> either (i) are imposed by legibility conditions, or (ii) fall out in some natural way from the computational process

THAT SECOND ONE IS WHAT I WAS JUST TALKING ABOUT. NOTE, INCIDENTALLY, THAT (i) AND (ii) ARE, IN A SENSE, REPRESENTATIONAL (i) AND DERIVATIONAL (ii). IS THERE ANYTHING IN THE NATURAL WORLD CORRESPONDING TO (i)? THAT'S THE TOUGH QUESTION WE MENTIONED BEFORE, CONCERNING THE NATURE OF A REPRESENTATIONAL SYSTEM. FOR WHAT IT'S WORTH, I WANT TO POINT OUT THAT THERE IS A CHARACTERISTIC 'STATIC' ASPECT TO (i), WHICH MIGHT ALLOW US TO SPEAK OF IT WITHOUT HAVING TO INVOKE TRICKY ISSUES ABOUT REPRESENTATIONALITY. THAT IS, (ii) IS CLEARLY DYNAMIC, IN THE CYCLIC SENSE WE'VE SEEN BEFORE, AND IN MORE WAYS THAT I'LL RETURN TO; IN CONTRAST, THE DEMANDS THAT INTERFACE SYSTEMS IMPOSE ON (i) MAKE THE RELEVANT RELATIONS RATHER FIXED. SIMILAR ISSUES ARISE IN ORGANISMS, WHICH DEBATE THEMSELVES BETWEEN STATIC AND DYNAMIC FORCES. WITHOUT DYNAMIC FORCES THERE WOULDN'T BE ANY METABOLISM, ANY REPRODUCTION, ANY DECAY; AT THE SAME TIME, WITHOUT STATIC POINTS

OF STABILITY THERE WOULDN'T BE ANY ORGANS, OR MORE BASICALLY SKELETONS, AXES OF VARIOUS SORTS, MEMBRANES... THERE ARE SOME SERIOUS CONJECTURES IN THE SYSTEMS LITERATURE AS TO HOW ONE MIGHT WANT TO CHARACTERIZE THIS DIALOGUE BETWEEN DYNAMIC AND STATIC PULSIONS. FOR INSTANCE, PEOPLE SPEAK OF DISSIPATIVE PHASE TRANSITIONS (E.G. GLOBAL CHANGES IN AN OPEN SYSTEM THAT, FOR INSTANCE, GOES INTO TURBULENCE) VS. CONSERVATIVE PHASE TRANSITIONS (E.G. GLOBAL CHANGES IN A QUASI-CLOSED SYSTEM, E.G. CRYSTALLIZATION). MY POINT IS THIS: WE MAY FIND GOOD METAPHORS AND EVEN WELL-STUDIED SYSTEMS IN THIS VIEW, WHICH ABSTRACTLY FITS RATHER NICELY WITH THE TWO PROPERTIES OF A GRAMMAR: THE STABILITY POINTS (STARTING WITH LEXICON ITEMS AND FINISHING WITH CHAINS) VS. THE DYNAMIC INTERACTIONS (MOVE, AGREE, AND SO FORTH). NOTICE THAT THE WORD 'SYMBOLIC REPRESENTATION' DOESN'T HAVE TO APPEAR HERE--WHETHER YOU NEED IT BEYOND THIS MACHINERY IS OF COURSE POSSIBLE.

{The significance of (ii) was brought out in Epstein's (1994) derivational analysis of c-command, which underlies much important work since, including Epstein et al. (1998).}<EPS = footnote>

Condition (A) requires that there is no Deep or Surface structure, or other levels that have been proposed.

WELL, I DON'T SEE THAT. I SEE THAT FOR SURFACE STRUCTURE, IN THAT IT IS AN *INTERNAL* SYSTEM--BUT D-STRUCTURE COULD BE AN INTERFACE WITH SOME OTHER SYSTEM. WE MIGHT NOT WANT THIS ON EMPIRICAL GROUNDS (KEARNEY'S ARGUMENT), BUT THAT'S A DIFFERENT STORY. IT DOESN'T *FOLLOW* FROM THE LOGIC OF THE PROPOSAL THAT D-STRUCTURE SHOULDN'T EXIST UNLESS WE PROVE THAT IT COULDN'T BE INTERFACING SOME REASONABLE SYSTEM, WHATEVER IT MIGHT BE. I'LL SAY MORE: THE WAY WE'VE LAYED THINGS OUT (WITH A NUMERATION, CONDITION (6) ON MERGE, THE INCLUSIVENESS/RECOVERABILITY CONDITION) IT WOULDN'T EVEN BE UNREASONABLE TO INTERPRET ALL THAT AS RESULTING FROM AN INTERFACE CONDITION WITH SOME SYSTEM--SAY HALE/KEYSER'S CONCEPTUAL STRUCTURE. AGAIN, THAT NEED NOT BE A *LEVEL OF REPRESENTATION* INTERFACE (AND COULD NOT BE ONE, IF KEARNEY IS RIGHT); BUT THERE COULD STILL BE A DYNAMIC CONNECTION WITH THAT CONCEPTUAL SYSTEM. IF SO, WHETHER WE CALL IT D-STRUCTURE OR WHETHER WE LIST ITS PROPERTIES SEPARATELY ((6), NUMERATION, ETC.) IS JUST A MATTER OF TERMINOLOGY. I SHOULD ADD HERE, INCIDENTALLY, THAT MANY WANT TO GET RID OF THE NUMERATION AND (6) FOR PRECISELY THAT REASON--SO THAT NO D-STRUCTURE RESIDUE REMAINS IN THE SYSTEM. THAT'S A NOBLE CAUSE, BUT ONE THAT, IT SEEMS, MISINTERPRETS THE STRONG THESIS. THE STRONG THESIS, TO REPEAT, JUST ASKS YOU TO OPTIMIZE THE INTERFACE CONNECTIONS. IF THE INTERFACES HAPPEN TO BE SIX HUNDRED AND SIXTY SIX, YOUR TASK IS NOT TO SEE HOW TO GET RID OF SIX HUNDRED AND SIXTY FOUR LEVELS, BUT RATHER TO FIGURE OUT A WAY OF OPTIMALLY CONNECTING WHAT YOU HAVE.

OF COURSE, IT'S EASIER TO CONNECT TWO THINGS THAN TO CONNECT THREE OR MORE, AS NEWTON SHOWED. BUT THAT'S A DIFFERENT POINT. NOW, YOU MIGHT WANT TO STOMP YOUR FOOT INSISTING ON THE EXISTENCE OF ONLY TWO INTERFACES--THE INEFFABLE SOUND AND MEANING. THAT, HOWEVER, SEEMS AN A PRIORI, THUS UNSCIENTIFICLY DOGMATIC POSTURE. FOR ALL I KNOW OR CARE, THERE MAY BE INTERFACES WITH THE MUSICAL SYSTEM. SURELY, A *WORKING* HYPOTHESIS COULD BE TO GO WITH TRADITION, AND SAY 'WELL, ALRIGHT, BUT SINCE WE DON'T KNOW HOW TO TEASE THE DIFFERENT ASPECTS OF MEANING, WE'LL LUMP THEM ALL TOGETHER INTO AN INTERFACE.' THAT'S FINE, BUT THEN YOU HAVE TO BE PREPARED TO SEE THE VARIOUS POSSIBILITIES, IF THEY MAKE SENSE, AND YOU SHOULDN'T REALLY BANK ALL YOUR MONEY ON GETTING RID OF A WELL-ESTABLISHED COMPONENT OF THE SYSTEM (E.G. (6)) JUST BECAUSE IT WOULD BE MORE MINIMALIST OTHERWISE. IT ISN'T. THE LOGIC OF 'LESS LEVELS IS BETTER THAN MORE LEVELS' ONLY WORKS *UP TO FACTUALITY*. TRIVIALY, THUS, FOR INSTANCE, A SINGLE LEVEL OF, SAY, LF WOULD BE EVEN BETTER, RIGHT? EXCEPT WE WOULDN'T HAVE HUMAN LANGUAGE, WE'D HAVE TELEPATHY. THIS IS RATHER SERIOUS, AND YOU SHOULD BE AWARE OF WHAT YOU'RE SAYING WHEN YOU INSIST ON ONLY LF AND PF *BECAUSE IT'S SIMPLER* THAN THAT PLUS D-STRUCTURE. YOU SHOULD REALLY ASK YOURSELF WHETHER IT GIVES YOU BOTH 'CONCEPTUAL NECESSITY' AND, OF COURSE, THE FACTS--FOR INSTANCE THE FACTS THAT (6) IS MEANT TO COVER. WE'LL RETURN TO ALL THIS LATER ONE IN THE COURSE.

It holds that everything accounted for in these terms has been misdescribed and is better understood in terms of legibility conditions at the interface: that includes the Projection Principle, Binding Theory, Case theory, the Chain Condition, etc.

Condition (B) is transparently false.

RIGHT, ALTHOUGH WE'LL HAVE TO ASK SOME TOUGH QUESTIONS HERE, LIKE WHETHER THE FEATURES THAT DO NOT FIT THE INTERPRETABILITY DESIGN ARE REALLY PART OF THE CORE LEXICAL ITEM, OR THEY ARE BORN AT THE NUMERATION (THAT WOULD MAKE ALL THE MORE SENSE IF IT IS THERE THAT A CONNECTION WITH THE CONCEPTUAL SYSTEM IS CHECKED).

Condition (C) permits rearrangement of lexical items LI and of elements constructed in the course of derivation, and deletion of features of LI. But optimally, nothing more.

AND EVEN THAT MIGHT NEED TO BE STREAMLINED...

Condition (D) has to be spelled out. Properties induced by legibility conditions might include adjacency, argument structure, scope, etc. Those of category (ii) should include at least (perhaps at most) the relations provided directly by the indispensable computational operation Merge. But there should be no government, no stipulated properties of chains, no binding relations internal to language, no interactions of other kinds.

THIS, OF COURSE, IS THE CENTRAL PHILOSOPHY OF THE PROGRAM. THE ONLY NOTE OF CAUTION I'D ADD IS THE SAME ONE I'VE BEEN PRESSING. IT'S AN EMPIRICAL ISSUE WHETHER, SAY, CHAINS HAVE A SEPARATE LEVEL OF INTERFACE. I DOUBT THAT'S THE CASE, BUT MY REASONS ARE EMPIRICAL. OR PUT DIFFERENTLY, YOU COULD IMAGINE MARTIANS HAVING FL LIKE WE DO AND A SYSTEMATIC INTERFACE WITH A CHAIN SYSTEM, IN WHICH CASE IT WOULD BE JUSTIFIED TO HAVE A CHAIN COMPONENT WHERE YOU WOULD HAVE OPERATIONAL PROPERTIES OF CHAINS IN SOME FORM; IF HOWEVER THIS IS NOT FL, THEN THE PROPERTIES OF CHAINS IN QUESTION (UNIFORMITY, COMMAND, COLLAPSE AT SOME POSITION, AND SO ON) SHOULD FOLLOW FROM SYSTEMIC DYNAMICS. THAT'S WHERE THE GAME BECOMES REALLY INTERESTING.

It is hardly necessary to observe that all of this is highly unlikely.

HERE I FEEL CHOMSKY IS BEING BOTH TOO CAUTIOUS AND TOO MODEST. THESE ARE SURELY UNLIKELY SCENARIOS, BUT SO IS EVERYTHING ELSE IN MIND AND, INDEED, EVEN IN LIFE. HEY, LIFE IS AN UNLIKELY EVENT, IT SEEMS! NONETHELESS, IF YOU'RE INTERESTED IN THE STUDY OF THE EMERGENCE OF FORM IN THIS UNIVERSE (THE BASIS FOR LIFE, MIND, AND OTHER SYSTEMS), THE LOGIC OF CHOMSKY'S PROGRAM IS NOT ONLY HIGHLY LIKELY, BUT INDEED THE ONLY GAME IN TOWN (WHICH OF COURSE DOESN'T MEAN IS RIGHT).

There is substantial empirical evidence supporting the opposite conclusion at every point.

THAT *IS* A FAIR POINT. PARTICULARLY BECAUSE:

Furthermore, a basic assumption of the work in the P&P framework, with its impressive achievements, is that everything just suggested is false: that language is highly "imperfect" in these respects.

HOWEVER, IT MIGHT BE TIME TO REFLECT ON WHY THAT WAS THE P&P LOGIC. (HERE I'LL BE DOING HISTORIC RECONSTRUCTION, HENCE I'LL BE EVEN MORE PERSONAL THAN BEFORE, SO TAKE EVERYTHING I SAY WITH A TON OF SALT.) I SEE TWO LINES TO EXPLORE, FOR THE 'HISTORIAN OF MODERN LINGUISTICS'... ONE: CHOMSKY CAME INTO THE ARENA OF MIND STUDIES WITH A BOMB FROM THE PAST: AS SOME NATURAL PHILOSOPHERS HAD BELIEVED, FUNDAMENTAL ASPECTS OF LANGUAGE ARE INNATE. FROM THE MID FIFTIES UP TO THE MID EIGHTIES THIS ONLY HAD ONE INTERPRETATION WITHIN MAIN-STREAM BIOLOGY, PSYCHOLOGY, AND SO ON: WHAT PINKER CALLS "THE LANGUAGE INSTINCT". ACTUALLY, THERE WERE TWO SUB-INTERPRETATIONS. ONE, ADVOCATED BY PINKER, BICKERTON, JACKENDOFF, NEWMAYER, AND OTHERS, IS ESSENTIALLY ADAPTATIONIST. LANGUAGE FORMS EMERGED BECAUSE OF SOME FUNCTIONAL NEED. ANOTHER ONE, ADVOCATED BY FODOR, BERWICK, OR LIGHTFOOT IS WHAT YOU MAY CALL, FOLLOWING GOULD, 'EXAPTATIONIST'.

THAT IS, LANGUAGE FORMS DID NOT EMERGE BECAUSE OF SOME FUNCTIONAL NEED, BUT RATHER AS THE BY-PRODUCT OF SOME OTHER FORM, WHOSE ORIGIN MAY OR MAY NOT, ITSELF, BE AS AN ADAPTATION (E.G. THE BRAIN GREW LARGER FOR SOME REASON, AND THAT LIBERATED ENOUGH 'MIND SPACE' TO ALLOW FOR LINGUISTIC STRUCTURES). THE LATTER POSITION IS THE ONE THAT, SO IT HAS ALWAYS SEEMED, ACCORDS BEST WITH THE FACTS OF LANGUAGE (THE KEY ONE BEING THE AUTONOMY OF SYNTAX). BUT NOTE AN IMMEDIATE CONSEQUENCE IT HAS: SINCE YOU'RE TRYING TO SEE THE LANDMARKS OF A KIND OF INNATE INSTINCT, ONE FURTHERMORE ARISING FROM AN EXAPTATION (IN THE 'RIGHT' VIEW), YOU WELCOME ANY SUB-OPTIMAL QUIRK OF LINGUISTIC FUNCTION THAT YOU ENCOUNTER. FOR INSTANCE, YOU LOOK AT 'CENTER EMBEDDING' AND YOU SAY: 'SEE? THAT STUPID, INNATE LINGUISTIC FORM LEADS TO SOMETHING THAT IS UTTERLY UNCOMMUNICABLE.' APART FROM BEING A CHALLENGE TO THE ADAPTATIONIST VIEW, THAT WAS BIG-TIME ANTI BEHAVIORIST AND SO ON, FOR HOW COULD YOU EXPLAIN THE ACQUISITION FROM THE ENVIRONMENT OF A KIND OF FORM THAT DOESN'T MAKE ANY GOOD SENSE? NOW, CHOMSKY HAS ALWAYS BEEN A BIT DISTANT ON THIS DEBATE, MAKING CRYPTIC REMARKS HERE AND THERE ABOUT LANGUAGE BEING 'LIKE A CRYSTAL' OR REMISCENT OF 'FIBONACCI PATTERNS', USUALLY NOT IN PRINT. ONE INTERPRETATION OF THIS DISTANCE IS THAT CHOMSKY WAS ACTUALLY BEHIND THE 'RIGHT' VIEW, ALBEIT WITH A TWIST: LANGUAGE ISN'T STRUCTURALLY SUB-OPTIMAL. NOTICE, BY THE WAY, THAT THIS DOESN'T AFFECT AT ALL THE LOGIC OF THE 'RIGHT' VIEW, WHICH ONLY DEMANDS SUB-OPTIMALITY *OF FUNCTION* AND HAS ABSOLUTELY NO STAKE ON WHETHER THE FORMS HAPPEN TO BE OPTIMAL OR NOT. THESE TWO, HOWEVER, AREN'T ALWAYS EASY TO SEPARATE, FOR SOMETIMES THE OPTIONAL FORM DOES HAVE AN ASSOCIATED OPTIMALITY OF FUNCTION (E.G. IF DERIVATIONS ARE COMPUTATIONALLY OPTIMAL, THAT WILL NO DOUBT AID THE PARSING OF CORRESPONDING REPRESENTATIONS). CONSEQUENTLY, SOMETIMES LACK OF OPTIMALITY IN STRUCTURE WAS SEEN AS A WELCOME TURN OF EVENTS IN THE DEBATE I'VE JUST MENTIONED (FOR THE 'RIGHT' VIEW ANYWAY), SINCE IT WAS MORE KINDLING TO THE NATIVIST ANTI-ADAPTATIONIST FIRE.

BUT THERE IS A SECOND INTELLECTUAL LINE WHOSE PROPERTIES HAVE, I THINK, AFFECTED CHOMSKY'S GEARING OF THE P&P SYSTEM IN A DIFFERENT DIRECTION. IN RETROSPECT (AND ASIDE FROM THE TREMENDOUS WEALTH OF DATA THAT THE MODEL HAS UNEARTHED), THERE WERE TWO ARCHITECTURAL LINES THAT HAVE POWERED P&P RESEARCH IN THE EIGHTIES, AFFECTING MUCH OF CHOMSKY'S OWN RESEARCH. THESE ARE THE EXTENDED PROJECTS OF KAYNE AND LASNIK; BY THAT I MEAN NOT JUST THEIR WORK WITH THEIR VARIOUS ASSOCIATES, BUT ALSO THE REACTIONS FROM MANY OTHER COLLEAGUES. KAYNE'S PROJECT WAS FOUNDED ON THE MATHEMATICAL PROPERTIES OF TREES AND THE VARIOUS INTERACTIONS THEY DO AND DO NOT ALLOW. LASNIK'S, ON THE COMPUTATIONAL DYNAMICS OF THE TRANSFORMATIONAL COMPONENT AND ITS VARIOUS RESULTS. ALL THE CONNECTEDNESS AND LCA LITERATURE IS BASED ON THE FIRST LINE OF

RESEARCH, WHEREAS ALL THE EXTRACTION AND BINDING LITERATURE IS BASED ON THE SECOND. ASIDE FROM ALL THE EMPIRICAL RESULTS, TWO PROPERTIES IMMEDIATELY JUMP OUT OF THESE LINES: THEY ENRICHED THE SYSTEM WITH NEW NOTATIONS (PATHS, SUBSTANTIVE PROJECTION LEVELS, ABSTRACT FEATURES, INDICES) AND THEY ALLOWED FOR VERY RICH INTERACTIONS, BE THEY AT THE TREE LEVEL (CONNECTEDNESS) OR AT THE DERIVATIONAL LEVEL (AFFECT @). IT IS IMPORTANT TO SEE THAT THESE SYSTEMS WERE CONCEPTUALLY SPOTLESS, THE LOGICAL EXTENSIONS OF (AND TO SOME EXTENT BASIS FOR) THE GOVERNMENT AND BINDING THEORY OF LGB. THE QUESTION, THOUGH, WAS THIS: DO WE WANT A SYSTEM WITH SUCH NOTATIONAL AND OPERATIONAL RICHES? MY SENSE IS THAT CHOMSKY ATTEMPTED THE MINIMALIST MOVE IN THE LATE EIGHTIES WITHIN THE CONTEXT OF ADDRESSING THAT INTERNAL QUESTION.

IT IS CURIOUS, HOWEVER, THAT THESE TWO LINES CONVERGE, UNDER CERTAIN ASSUMPTIONS. OF COURSE, WE ALL STILL WANT TO BE NATIVIST AND MANY OF US STILL HAVE NO REASON TO GIVE UP THE 'EXAPTATIONIST' LINE (WHERE FORM DOESN'T GENERALLY FOLLOW FROM FUNCTION). BUT DO WE HAVE ANYTHING ABOUT THE EMERGENCE OF LINGUISTIC FORM OTHER THAN ITS NOT BEING AN ADAPTATION? HERE CHOMSKY'S OLD METAPHOR OF THE CRYSTAL AND THE FIBONACCI PATTERNS MAY WELL FIND A CURRENT INSTANTIATION IN A SYSTEM THAT AVOIDS THE OPERATIONAL AND NOTATIONAL EXCESSES OF THE PAST, HENCE (ARGUABLY) REALISTICALLY IN TERMS OF PRESENT-DAY MIND STUDIES THAT MAY ASSUME D'ARCY THOMPSON'S GENERAL QUESTION FOR BODY PLANS.

It would be no small task, then, to show that this apparatus is a kind of descriptive technology, and that if we abandon it, we can maintain or even extend descriptive and explanatory force.

I READ THIS SENTENCE IN THE TERMS EXPRESSED IMMEDIATELY ABOVE, AS I DO THE NEXT. THERE'S A SENSE OF EXCITEMENT IN THE NEXT SENTENCE, ACTUALLY, WHICH I FEEL PERTAINS NOT JUST TO THE PROGRESS WE'VE WITNESSED IN LINGUISTICS, BUT MORE GENERALLY IT SEEMS TO ME ALLUDES, ALSO, TO PROGRESS IN OTHER (MORE OR LESS RELATED) FIELDS. IN MY VIEW, THOUGH STILL CONTROVERSIAL IN MANY CAMPS, IT IS NO LONGER WEIRD TO ARGUE FOR THE INNATENESS OF LANGUAGE. THE ISSUES NOW ARE MORE SUBTLE, PERTAINING TO PROFOUND QUESTIONS ABOUT FORM WHICH ARE, NO DOUBT, STILL HIGHLY NON-STANDARD. BUT (PERHAPS OPTIMISTICALLY) I FEEL AS IF WE'VE MOVED ONE NOTCH FORWARD IN STUDIES OF COMPLEX SYSTEMS, AND (BY FAR) WE DON'T KNOW ANY COMPLEX SYSTEM AS WELL AS WE DO THE LINGUISTIC ONE.

Nevertheless, recent work suggests that such conclusions, which seemed out of the question a few years ago, are at least plausible, possibly correct in nontrivial respects.

Much of <MP> is devoted to establishing parts of (30) in terms of bare phrase structure, with a

cyclic notion of generalized transformations and reinterpretations of Binding and Case theory. I will assume that these are the right directions, though many serious problems remain.

Consider Condition (C). It requires that there be no phrasal categories or bar levels, hence no X-bar theory or other theory of phrase structure, apart from bare phrase structure, which will be simplified further below. It also rules out introduction of traces, indices, lambda operators, and other new elements in the course of operation of C<HL>. Recourse to such devices could be innocuous (e.g., if used for convenience to annotate properties that can be determined by inspection at LF), but questions arise if they enter into interpretation and function significantly within the computation -- for example, percolation of indices, operations that apply specifically to trace, etc.

THIS IS ALL STRAIGHTFORWARD, AND IS WHAT THE CONSERVATION LAW IS SUPPOSED TO PREVENT. INCIDENTALLY, YOU CAN THINK OF THESE SORTS OF QUESTIONS IN OTHER DOMAINS AS WELL. SO IMAGINE YOU WERE STUDYING PROTEIN FOLDING AND YOU NEEDED INDICES TO CODE A DEPENDENCY BETWEEN TWO LONG-DISTANCE CHUNKS OF PROTEIN. WOULD YOU TAKE THE INDICES TO BE REAL? PERHAPS, BUT NOT WITHOUT CONSEQUENCE, SURELY. ALTERNATIVELY, THE INDICES ARE JUST MACHINERY WITHOUT A BIO-CHEMICAL BASE. MUTATIS MUTANDIS, WE MAY BE RIGHT OR WRONG ABOUT THE EXISTENCE OF INDICES, BAR LEVELS, OR EVEN SYMBOLS THEMSELVES. BUT THE STRONG THESIS FORCES YOU TO CONSIDER THE CONSEQUENCES OF ANY OF THE DECISIONS YOU MAKE HERE.

In <MP>, indices are introduced, but not within C<HL> itself; rather, in extending lexical arrays to numerations (see note <NUMER>). This narrow departure from Condition (C) could be eliminated by weakening the requirement that an item of a lexical array be removed when accessed in computation, leaving this as an option.

IF SO, THE NUMERATION WOULD IN ESSENCE NOT BE A SET OF TOKENS, BUT A SET OF TYPES, AS FAR AS I CAN SEE. THE ISSUE IS CONFOUNDED BY A RELATED QUESTION: IS THE LEXICON A SET OF TYPES OR A SET OF TOKENS? THIS SEEMS LIKE A MORE TRIVIAL QUESTION THAN IT IS. IF YOU'RE TEMPTED TO THINKING THAT OF COURSE THE LEXICON IS A SET OF TYPES, TELL ME WHAT'S YOUR ARGUMENT. I ONLY KNOW OF ONE, WHICH TRIES TO REDUCE THE ALTERNATIVE VIEW TO THE ABSURD--AND IT PROBABLY BEGS THE QUESTION. IF YOU'RE GOING TO SAY THAT THE LEXICON IS A SET OF TOKENS, YOU'LL NEED A COPYING MECHANISM, SO THAT EVERY TIME YOU ACCESS THE LEXICON YOU ACTUALLY COPY SOME TOKEN, AND USE IT IN A SENTENCE. IS THIS COPYING MECHANISM CONCEPTUALLY NECESSARY? WELL, ONE THING IS ARGUABLE: IT INTUITIVELY ADDS INFORMATION: PRIOR TO THE COPY YOU HAD ONE TOKEN, AFTER THE COPY YOU HAVE ANOTHER ONE (OR PERHAPS A TOKEN SEPARATED FROM THE LEXICON, IN THE SPACE OF A DERIVATION). THEN AGAIN, THE CONSERVATION LAW THE WAY IT IS USUALLY UNDERSTOOD RESTRICTS THE INPUT-OUTPUT RELATION IN THE DERIVATION, THAT IS, BETWEEN THE LEXICAL ARRAY AND THE LEVELS OF REPRESENTATION. HERE, THOUGH, WE'RE NOT

TALKING ABOUT DERIVATIONS, BUT ABOUT WHAT GOES ON *PRIOR* TO DERIVATIONS. CHOMSKY HAS OPENED THAT CAN OF WORMS WHEN TALKING ABOUT COMPLEXITY IN A DERIVATIONAL SIZE, HIS SOLUTION BEING IN TERMS OF REDUCING LEXICAL ACCESS. THE QUESTION IS WHETHER THAT PRE-DERIVATIONAL MAPPING (FROM THE LEXICON TO THE LEXICAL ARRAY) OBEYS SOME CONSERVATION LAW, AND IF SO WHAT FORMAT IT HAS. IF IT DOES, AND THE MAPPING CANNOT CREATE STRUCTURE, THEN YOU WON'T BE ABLE TO JUST HAVE TOKENS IN THE LEXICON, YOU'LL NEED TYPES TO INSTANTIATE. BUT THIS IS NOT MUCH BETTER, TO TELL YOU THE TRUTH, SINCE NOW YOU MUST REALISTICALLY ASK YOURSELF 'WHAT'S A TYPE?' REMEMBER, THE NOTION 'TYPE' WAS INVENTED BY ARISTOTLE AND WORKED OUT BY RUSSELL, BUT NOTHING IN ITS PHILOSOPHICAL HISTORY FORCES YOU TO ASSUME IT, OR WORSE STILL, TO TAKE IT FOR GRANTED. BLUNTLY PUT, IF TYPES EXIST IN MIND, YOU HAVE TO SEE WHAT THAT MEANS, AND WHAT IT MEANS FOR THEM TO DIFFER FROM THEIR INSTANTIATIONS INTO TOKENS. FOR INSTANCE, ARE TYPES PARADIGMATIC AND, IN SOME SENSE, IN LONG-TERM MEMORY, WHEREAS TOKENS OF THOSE TYPES ARE SYNTAGMATIC AND, AGAIN IN SOME SENSE, IN SHORT-TERM MEMORY? PERHAPS--WHO KNOWS. BUT THAT SORT OF THING IS WHAT'S AT STAKE, AT THE VERY LEAST.

Whether that is the correct move depends on the consequences. A minor matter is that it would require modification of interpretive procedures at the LF interface.

THAT IS, IF THE INTERFACE DOESN'T DEAL WITH IDENTIFYING TOKENS FROM THE NUMERATION, YOU'LL HAVE TO TELL THE INTERFACE WHAT PROCEDURE IT MUST FOLLOW IN ORDER TO GET 'ANOTHER ONE OF THOSE'. NOTHING TRIVIAL, FOR SURE. ONE MIGHT EVEN WONDER WHETHER SUCH THINGS AS THETA-ROLES OR CASE FEATURES (OR OTHER THINGS) AREN'T TELLING LF PRECISELY THAT. NOTE, INCIDENTALLY, THAT THAT SPECULATION WOULD ADDRESS ONE OF MY WORRIES FOR THE VERSION OF THE SYSTEM WHERE LF IS THE SINGLE 'SEMANTIC INTERFACE'. MY CONCERN WAS HOW TO TAKE THAT SERIOUSLY, WHEN LF IS A NON-CONFIGURATIONAL SYSTEM (A COLLECTION OF CHAINS) WHILE THETA STUFF IS CONFIGURATIONAL. BUT YOU MIGHT RESPOND TO ME THUS: WITHOUT THETA STUFF YOU WOULDN'T EVEN BE ABLE TO TELL APART THE DIFFERENT TOKENS THAT ENTER LF, IF YOU DON'T MARK THEM IN THE NUMERATION. THAT'S CONCEIVABLE, AND WE SHOULD RETURN TO IT LATER ON.

More important, it would require a new notion of <chain>. That seems to be the only significant consequence, and it merits some thought.

THAT'S PRECISELY THE POINT I WAS TRYING TO MAKE. I'LL POSTPONE FURTHER DISCUSSION OF THESE ISSUES UNTIL LATER, THOUGH.

INCIDENTALLY, ONE OTHER THING. THE INDICES IN ITEMS IN THE NUMERATION WERE NEVER REALLY NEEDED--AND THAT'S NOT THE PROBLEM. THAT IS, YOU

COULD SAY THE NUMERATION IS A SET OF TOKENS, AND CLEARLY A TOKEN OF A MAN IS NOT THE SAME AS ANOTHER TOKEN OF A MAN, OR THERE WOULDN'T BE YOU AND ME. BUT OF COURSE, WHETHER THE NUMERATION IS A SET OF TOKENS IS DEEP, AND WHAT WE'RE DISCUSSING. (IF IT IS, IS IT BECAUSE WE'RE COPYING TOKENS FROM AN ALL TOKEN LEXICON--WHICH I DOUBT--OR BECAUSE WE'RE SOMEHOW INSTANCIATING THOSE TOKENS FROM THE LEXICON INTO A SHORT-TERM-MEMORY TRACK--WHICH I SUSPECT--AND IF THE LATTER WHAT DOES IT MEAN TO INSTANCIATE SOMETHING? ALL FAMILIAR, TOUGH STUFF.)

If _ in the syntactic object SO is merged somewhere else (by the operation Move) to form SO', then the two occurrences of _ constitute a chain, the original occurrence called the "trace" or "copy" of the new one.

THIS IS THE TRADITIONAL TAKE ON CHAINS, OFTEN MISUNDERSTOOD.

The terminology is misleading, for several reasons. First, each of the elements is a "copy" of the other;

OF COURSE, MATTERS HERE GET EVEN MORE MESSY, SINCE THESE COPIES ARE INTERNAL TO THE SYSTEM, THUS NOT REALLY COPIES OF TOKENS, BUT (AS USED TO BE CALLED IN LSLT) OCCURRENCES OF THOSE TOKENS. NOW YOU HAVE TO WORRY ABOUT SOMETHING ELSE, AGAIN. YOU HAD TYPES AND TOKENS AND NOW YOU HAVE TOKENS AND OCCURRENCES. ARE THE RELATIONS AMONG EACH OF THOSE THE SAME? MATHEMATICALLY THAT COULD BE POSSIBLE, WITH OCCURRENCES BEING TOKENS OF TOKENS UNDERSTOOD AS TYPES. MY HUNCH, FOR WHAT IS WORTH, IS THAT REALISTICALLY THIS IS NOT WHAT'S HAPPENING AT ALL, AND IF WE'RE TO JUDGE FROM THE LONG-TERM/SHORT-TERM DISTINCTION IN TERMS OF ACTIVATION OR THE PARADIGMATIC/SYNTAGMATIC CHARACTERIZATION IN TERMS OF THEIR FORM, OCCURRENCES ARE NOT TOKENS AT ALL; IN FACT, THEY ARE NOT EVEN 'MATERIAL' IN A SENSE THAT I RETURN TO--THEY ARE MERE OPERATIONAL CONSTRUCTS THAT SURVIVE ONLY WHILE THE DERIVATION IS ACTIVE. MY STRONGEST ARGUMENT FOR THIS IS IN TERMS OF THE PARADIGMATIC/SYNTAGMATIC CUTS. THAT IS, THE LEXICON IS REAL IN A PARADIGMATIC SENSE. WE WITNESS LEXICAL RULES, FOR INSTANCE, IN TERMS OF THE PANINIAN ELSEWHERE CONDITION, OR CANONICITY RESTRICTIONS, WHICH HOLD OF PARADIGMS. THERE IS NO SUCH THING INSIDE THE SYNTACTIC ENGINE, JUST AS--I'LL TRY TO ARGUE, CONTRA GENERATIVE SEMANTICISTS--THERE IS NO SYSTEMATICITY, TRANSPARENCY, AND REGULARITY IN THE LEXICON, ALL SYNTAGMATIC PROPERTIES. IF YOU USE THIS AS YOUR WEDGE TO SEPARATE TYPES FROM TOKENS, THEN IT IS PLAINLY THE CASE THAT THERE ARE NO PARADIGMATIC PROPERTIES THAT HOLD INTERNAL TO THE MEMBERS OF A CHAIN, HENCE NO JUSTIFICATION IN THINKING IN THEM AS TYPES OF ANYTHING BY MY CRITERIA. INDEED, I SUSPECT THAT *NO* PROPERTIES HOLD INTERNAL TO THE MEMBERS OF A CHAIN, WHICH IS WHY I THINK IT IS A

MISTAKE TO THINK OF THEM AS WHAT I'M CALLING 'MATERIAL'.

second, "copy theory" is the simplest version of transformational grammar, making use only of Merge, not Merge followed by an operation that deletes the original; and under trace theory, a further operation that creates a new kind of element, trace, a serious violation of inclusiveness. These are "imperfections," to be avoided unless shown to be necessary. I will continue to use the terminology, but only for expository convenience, adopting the "copy theory" as the null hypothesis.

THESE ARE ALL OPERATIONAL POINTS, BUT FAIRLY REASONABLE ONES.

{That the "copy theory" is the simplest version is clear in the earliest formulation, in Chomsky (1955-6). The more complex operation Merge-Delete was adopted there on the assumption that T-markers are mapped to PF. When that picture was modified under EST, Delete was abandoned in favor of trace theory, trace being a new kind of element. That was an error (mine), and the "copy theory," which restored the simplest case of the original approach, was mistakenly regarded as a further innovation. The divergence of history and logic has caused much confusion.}

THIS IS A VERY INTERESTING FOOTNOTE, HISTORICALLY. IT IS WORTH KEEPING IN MIND, BUT IT IS ALSO IMPORTANT TO UNDERSTAND THAT WHETHER COPY THEORY IS SIMPLER OR NOT THAN MOVE--AND I DO THINK IT IS--DEPENDS ON ADDRESSING THE COMPLEX MORASS OF QUESTIONS JUST POSED.

If LAs are extended to numerations NUM, a chain can be defined as a sequence of identical _'s; more accurately, a sequence of <occurrences> of a single _.

THIS IS WHAT I WAS SAYING. THESE ARE NOT TOKENS OF _, REMEMBER. IF YOU HAVE A HARD TIME UNDERSTANDING HOW A COPY OF _ IS NOT A SEPARATE TOKEN, YOU'RE PERHAPS BEING A BIT TOO LITERAL WITH HOW THE MATERIAL UNIVERSE OPERATES. TAKE THE FAMOUS 'TRANSPORTER' IN STAR TREK. IT'S SUPPOSED TO WORK BY MEANS OF 'TELEPORTATION' (SOMETHING WHICH, INCIDENTALLY, HAS BEEN RECENTLY ACHIEVED FOR A LASER BEAM, I BELIEVE). IT GOES LIKE THIS: YOU'RE SOME SET OF PARTICLES, AND THE TRICK IS TO FIND (BY SOME MEANS) AN IDENTICAL SET OF PARTICLES IN ANOTHER REGION OF THE UNIVERSE. BY THE WAY, YOU'RE PROBABLY ALREADY TROUBLED: HOW CAN THE TWO SETS BE IDENTICAL IF THEY ARE IN DIFFERENT REGIONS? WELL, CHECK YOUR INTUITIONS ABOUT IDENTITY: HOW DO YOU KNOW THAT TWO ELECTRONS ARE *NOT* IDENTICAL? WHEN THINGS GET AT THAT LEVEL, THEY END UP BEING SO SIMILAR THAT, EXCEPT FOR TIME/SPACE (OR SIMILAR DIMENSIONS, OFTEN MORE ARCAIC ONES) THEY ARE QUITE LITERALLY THE SAME BY MOST CRITERIA. CERTAINLY, YOUR PARTICLES ARE CHANGING CONSTANTLY, AND THE FACT THAT THEY ARE DIFFERENT (BY SOME CRITERION) WHEN YOU'RE BORN AND WHEN YOU DIE DOESN'T MEAN THERE ARE TWO DIFFERENT YOU'S OUT THERE. SO THEN THE TRICK IS THAT, FOR ALL THAT COUNTS (E.G. CONSCIOUSNESS AND MEMORIES IN STAR TREK...), THE TWO

SETS OF PARTICLES (YOURS AND THE OTHER) ARE INDEED IDENTICAL. AT THAT POINT, SO THE STORY GOES, YOU'VE BEEN TELEPORTED. OF COURSE, THE TRICK IS TO FIND OUT THE EXACT SAME PARTICLES AND TO ARRANGE THEM APPROPRIATELY; HOWEVER, THIS IS TRICKY ONLY IN THEORY: YOU'RE VERY COMPLEX. IT APPARENTLY IS NOT SO TRICKY ANYMORE FOR WELL BEHAVED CHUNKS, LIKE THE BEAM I'VE JUST MENTIONED. A SECOND PROBLEM IS ALSO OBVIOUS: WHY DOESN'T THE PROCEDURE CREATE TWO YOU'S? AS IT TURNS OUT, FOR PROPERTIES OF QUANTUM REALITY THAT I DON'T CLAIM TO UNDERSTAND, YOU CANNOT MAKE ONE OF THESE 'COPIES' AND KEEP THE ORIGINAL--SOMETHING IN THE EQUATIONS FAILS IF YOU DO THAT. THE WAY I VISUALIZE IT IS THUS: YOU STRETCH AND BOND AND KNOT A GIVEN PART OF THE UNIVERSE TO GET YOUR PATTERN OVER THERE, AND AS A CONSEQUENCE OF ALL THAT TWISTING AND SHOUTING ANOTHER PART OF THE UNIVERSE (THE PRESENT YOU) DISENTANGLES. ANYWAY, THAT'S TELEPORTATION, AND REMEMBER, THE ISSUE WAS: DID WE CREATE A SEPARATE TOKEN OF YOU WHEN WE TELEPORTED YOU? WELL, NOT COMPLETELY OBVIOUS, AT LEAST IN STAR TREK (ALTHOUGH A GLORIOUS EPISODE PLAYS WITH PRECISELY THIS ISSUE, HAVING RYKER DUPLICATED AND APPEARING LATER ON TO CONFRONT THE 'REAL' RYKER--SOMETHING WHICH, AS I SAY, IS APPARENTLY PHYSICALLY IMPOSSIBLE). MORE TO THE POINT: CONCEPTUALIZE THE MOMENT IN WHICH WE ACTUALLY ARE REPRODUCING YOU OUT THERE, SO THAT WE SORT OF SIMULTANEOUSLY (ALTHOUGH THAT'S REALLY NOT THE RIGHT NOTION) HAVE YOU HERE AND THERE. ARE YOU THEN TWO TOKENS OF YOU? OR IS THAT A MERE ARTIFACT ABOUT HOW THE SYSTEM WORKS? TRICKY ISSUES--EVEN IN PHYSICS! AT ANY RATE, I SUGGEST THAT WHEN YOU EXPLORE THESE MATTERS PERTAINING TO COPIES, CHAINS, AND SO ON, YOU GO WITH WHAT WORKS, ON A MATHEMATICAL SENSE, FORGETTING ABOUT MORE OR LESS NAIVE INTUITIONS ABOUT COPYING A MATERIAL THING, OBJECTS HAVING TO BE IN A SINGLE POSITION, AND OTHERWISE TAKING THE COPIED OBJECTS, AS IT WERE, TOO SERIOUSLY. I MEAN, TAKE THEM SERIOUSLY AS FORMAL OBJECTS, BUT DON'T CONFUSE THE METAPHOR OF AN OPERATIONAL COPY WITH THE CREATION OF AN OBJECT IN THE SYSTEM.

That seems to be the simplest way to characterize chains, hence to be adopted unless shown to be empirically inadequate. Thus in (12iii), repeated as (31), there are two occurrences of $\langle I \rangle$ and two occurrences of $\langle a \text{ proof} \rangle$ (where the terms are spelled out and at the trace):

THAT IS, 'TERMS' IN THE SENSE OF THE $\{a, \{a, b\}\}$ NOTATION; IN OTHER WORDS, A CHAIN HAS, BY DEFINITION, MORE THAN ONE TERM (WE'LL RETURN TO THIS).

(31) $I_{\langle i \rangle} T [\langle t \rangle \langle \{i \rangle \rangle \text{ expect } [[a \text{ proof}]_{\langle j \rangle} \rangle \text{ to be discovered } \langle t \rangle \langle \{j \rangle \rangle]$

Chains so-defined can be formed only by movement, given $LA = NUM$.

THE CONSERVATION LAW AGAIN. OF COURSE, THERE IS AN EMPIRICAL ISSUE HERE; WHETHER, FOR INSTANCE, ELLIPSIS SHOULD BE CONSIDERED A CHAIN

PHENOMENON (EVEN IF IT OBVIOUSLY DOESN'T INVOLVE MOVEMENT, ETC.). WE'RE SETTING THAT ASIDE NOW, ASSUMING THE ANSWER IS NO.

Basic properties of chains should then follow from elementary derivational principles. That would be a good result, eliminating stipulated properties of chains and explaining why these hold.

IN FACT, IT WOULD BE MORE THAN JUST A GOOD RESULT; IT'S A NECESSARY RESULT IF THE SYSTEM IS DESIGNED THE WAY WE HAVE DESIGNED IT, WITHOUT AN INTERFACE LEVEL SPECIFICALLY FOR CHAIN PROPERTIES.

If achievable, it would also bear on broader questions that arose in sec. 3, providing support for the derivational interpretation of the recursive function $C\langle HL \rangle$, hence for the surprising conclusion that there is even an empirical issue concerning alternative interpretations.

THIS TOO IS VERY IMPORTANT (ALTHOUGH IT IS CONCEIVABLE, THOUGH I HAVEN'T SEEN THIS PROVEN) THAT NATURAL CHAIN PROPERTIES COULD ALSO BE OBTAINED IN A REPRESENTATIONAL FORMAL SYSTEM, AS A CONSEQUENCE OF MERE FORMAL INTERACTIONS.

What exactly do we mean by an "occurrence of _"? To ensure that occurrences in the intended sense are distinguished, we can take an occurrence of _ in K to be the full context of _ in K.

NOTE, THIS ISN'T JUST A METHODOLOGICAL ISSUE (HOW TO REFER TO THE DIFFERENT COPIES). GO BACK TO THE TELEPORTATION METAPHOR. IT WAS IMPORTANT TO TELEPORT YOU BOTH TO TRACK YOU DOWN (WHAT IS CALLED IN STAR TREK 'LOCKING INTO YOUR COORDINATES'...) AND TO TRACK DOWN THE REGION OF THE UNIVERSE WHERE YOU ARE TO REAPPEAR, INCLUDING THE GENEROUS PARTICLES OVER THERE THAT WILL RECONSTITUTE YOU (FORTUNATELY, THE UNIVERSE HAS A GOOD SUPPLY OF THE PARTICLES THAT MAKE YOU). WITHOUT KNOWING THE CONTEXT WHERE YOU ARE AND WHERE YOU WILL BE, THERE SIMPLY WON'T BE ANY TELEPORTATION FOR YOU, SINCE IN A SENSE THE TRICK IS TO TURN THE CONTEXT INTO YOU, WHILE UNPLUGGING YOU IN THE CONTEXT YOU INITIALLY OCCUPY (LETTING YOUR INITIAL PARTICLES GO FREE). IN OUR CHAINS TOO, YOU MUST KNOW THE REACH OF THE CHAIN, THE CONTEXT WHERE IT STARTS AND THE CONTEXT WHERE IT REAPPEARS. INDEED, IT IS PERHAPS BEST TO THINK OF THE CHAIN AS LINKING THOSE TWO CONTEXTS, JUST AS YOU CAN BE THOUGHT, UPON A TELEPORTATION, AS LINKING TWO GIVEN PHYSICAL CONTEXTS. IN A VERY LITERAL SENSE, THEN, A CHAIN IS A SORT OF STRETCHING OF A PHRASAL CONTEXT BEYOND THE NORMAL LIMITS IMPOSED BY THE MERGE OPERATION.

{We could, for example, identify this as $K' = K$ with the occurrence of _ in question replaced by some designated element OCC distinct from anything in K. In Chomsky (1955-6), "occurrence" is defined in terms of linear order, adapting a device from Quine's <Mathematical Logic>; but that mechanism is not available here.}

THIS IS A FORMAL FOOTNOTE TO TELL YOU HOW OCCURRENCES WERE CODED INITIALLY. FOR WHAT IT'S WORTH, I WOULDN'T BE TOO CAVALIER ABOUT DISMISSING QUINE'S NOTATION AS MERELY FORMAL. IT IS NOT ALTOGETHER CLEAR THAT MANY OF QUINE'S ARTIFACTS CAN BE REPLACED BY ANYTHING WITHIN THE CONFINES OF SET THEORY. FOR INSTANCE, TAKE HIS DEFINITION OF ORDERED PAIR $\{\{a\}, \{a, b\}\}$. WHAT ELSE CAN YOU DO? NOT MUCH. YOU COULD TRY $\{a, \{a, b\}\}$, IF YOU DON'T CARE ABOUT NOT ASSUMING THE FOUNDATION AXIOM (BUT THERE ARE REASONS FOR ASSUMING IT). YOU CANNOT JUST SAY 'I'LL PAINT ONE ELEMENT BLUE', SINCE THAT'S NOT A PREDICATE THAT SET THEORY ALLOWS WITHOUT APPROPRIATELY ENRICHING IT. SO IT MAY TURN OUT THAT YOU'RE STUCK WITH QUINE'S NOTATION FOR GOOD FORMAL REASONS. IF SO, THIS IS AS CONCEPTUALLY NECESSARY AS A SIMILAR POINT ABOUT DERIVATIONS. THE POINT IS, IN ORDER TO ARGUE FOR THE DERIVATIONAL SYSTEM YOU MUST BE FAIR WITH THE REPRESENTATIONAL ONE: THAT SYSTEM TOO MAY HAVE SOUND FORMAL LIMITATIONS THAT YOU'RE FORCED INTO FOR NON-TRIVIAL REASONS.

In $\langle MP \rangle$, a simpler notion is proposed: an occurrence of $_$ is a sister of $_$.

THIS IS SIMPLE, BUT IT MAKES A FANTASTICALLY INTERESTING CLAIM, ALONG THE LINES OF WHAT I WAS SUGGESTING ABOVE FOR QUANTUM REALITY. NOTE, A SISTER OF $_$ IS QUITE LITERALLY $_$ 'S CONTEXT (EXPLICITLY SO IN MP). SO TRANSLATING THE STATEMENT ABOVE, AN OCCURRENCE OF $_$ IS $_$ 'S CONTEXT. AS WE SAW, THAT'S A WAY OF DESCRIBING WHAT GOES ON IN YOUR TELEPORTATION, AND IT WOULD BE APPROPRIATE TO DEFINE EACH OF YOUR OCCURRENCES AS YOUR CONTEXT--INDEED, THAT'S THE *ONLY* WAY OF DEFINING YOUR OCCURRENCES (THE STARTREK COORDINATES).

Then in (31) the occurrence of matrix subject $\langle I \rangle$ is $I\langle 1 \rangle = \text{"T...discovered } \langle t \rangle \langle \{j\} \rangle \text{"}$ and the occurrence of its embedded copy is $I\langle 2 \rangle = \text{"expect...discovered } \langle t \rangle \langle \{j\} \rangle \text{"}$ (actually, the syntactic objects corresponding to them).

THAT IS, OBJECTS DEFINED AS $\{T, \{T, \dots\}\}$ AND $\{\text{EXPECT}, \{\text{EXPECT}, \dots\}\}$

The occurrences of $\langle a \text{ proof} \rangle$ and its trace are $P\langle 1 \rangle = \text{"to be discovered } \langle t \rangle \langle \{j\} \rangle \text{"}$ and $P\langle 2 \rangle = \text{"discovered,"}$ respectively.

WE SHOULD EMPHASIZE THIS. "discovered" IS ACTUALLY, AFTER MOVEMENT, AN OCCURRENCE OF THE CHAIN OF a proof!

In (31) two chains are defined: $C\langle I \rangle = \langle I\langle 1 \rangle, I\langle 2 \rangle \rangle$ and $C\langle P \rangle = \langle P\langle 1 \rangle, P\langle 2 \rangle \rangle$.

THIS IS CLASSICAL NOTATION, NOW SUBSUMED UNDER THE MORE INTERESTING IDEA ABOVE WHICH DOESN'T USE INDICES.

In informal description $C\langle I \rangle$ is the chain $\langle\langle I \rangle, \langle t \rangle \langle \{i\} \rangle\rangle$ and $C\langle P \rangle$ the chain $\langle\langle a \text{ proof} \rangle, \langle t \rangle \langle \{j\} \rangle\rangle$, where $\langle I \rangle$ and $\langle a \text{ proof} \rangle$ are called the "heads" of the respective chains.

IT IS WORTH ASKING WHETHER THIS TRADITIONAL NAME MAKES ANY SENSE. IN PF IT DOES, TO THE EXTENT THAT THE HEAD OF THE CHAIN IS WHAT GETS LEXICALIZED. IS IT A DEEP FACT ABOUT CHAIN CONSTRUCTION? PERHAPS HEADS OF CHAINS LEXICALIZE IN THE SAME WAY THAT HEADS OF PROJECTIONS (LEXICAL ITEMS) DO. PERHAPS CHAINS ARE, IN SOME SENSE, SUPER-PROJECTIONS WHICH TAKE AS OPERATIONAL TERMS NOT THE CONSTITUENTS OF MERGE, BUT RATHER MORE ELABORATE CONSTRUCTS THAT, AS IT WERE, UNDERGO A KIND OF SUPER-MERGE. THAT IS, NOTE THAT THE HEAD OF A CHAIN (E.G. a proof) SERVES AS A KIND OF LABEL TO RELATE TWO DISTANT (NON SISTER) TERMS, E.G. "discovered" AND "to be discovered t". LIKE MERGE, THE RELEVANT RELATION IS ASYMMETRIC, NOT ONLY IN THAT AT MOST ONE 'LINK' OF THE CHAIN FORCES THE PF COLLAPSE IN THAT POSITION, BUT ALSO IN THE UNSPOKEN ASSUMPTION THAT INTERPRETATION OF THE CHAIN-HEAD PARTS IS IN A SINGLE PLACE FOR LF REASONS, EVEN IF SCATTERED (AS IN OPERATOR-VARIABLE CONDITIONS). I WILL NOT PURSUE ANY OF THIS YET (SEE THE END OF CHAPTER FIVE OF RHYME AND REASON--HENCEFORTH R&R--IF YOU CARE), BUT I WANT TO NOTE THAT IT IS CONSISTENT WITH THE FORMALISM AND THE GENERAL OUTLINE SO FAR.

In (31), occurrences are properly identified if taken to be sisters, but that might not always be the case.

{Suppose, for example, that the operation Move could raise the object of V to become the object of V', forming [V' OB]. Then if $V = V'$ (by virtue of V-raising), the sisters of OB would be identical.}

THIS EXAMPLE IS NOT COMPLETELY STRAIGHTFORWARD, FOR EVEN WHEN V RAISES TO V' IT IS NOT OBVIOUS THAT, THEN, $V = V'$ --BUT THE MATTER OF PRINCIPLE REMAINS.

The simplified definition in terms of sisterhood is based on the assumption that other properties of $C\langle HL \rangle$ guarantee that no problems arise. I will assume that to be true.

We can simplify chains from sequences to sets, relying on the fact that a "higher" occurrence of $_$ properly contains lower ones. Thus in $C\langle I \rangle$ of (31), $I\langle 1 \rangle$ properly contains $I\langle 2 \rangle$ ($\langle I \rangle$ c-commands its trace).

THE DEFINITION IN CHAPTER 4 RELIED ON ORDERED PAIR NOTATION. AS IT TURNS OUT, THAT IS GENERALLY NOT NEEDED, FOR PRECISELY THE REASON CHOMSKY JUST MENTIONS (IT IS GENERALLY POSSIBLE TO DISTINGUISH THE TWO CHAIN LINKS WITHOUT MENTIONING LINEAR ORDER). THAT'S NICE, FOR OTHERWISE ONE WOULD HAVE TO MOTIVATE WHY THE LINKS ARE ORDERED, HOW THAT'S CAPTURED, AND SO ON--WHEREAS THE SIMPLEST NOTATION

MAKES FEWER ASSUMPTIONS.

Chains formed by successive-cyclic movement fall into place directly.

THAT IS, ASSUMING THEY ARE REAL, WHICH IS NOT COMPLETELY OBVIOUS IN THE CASE OF A-MOVEMENT (PARTICULARLY IF, AS CHOMSKY CLAIMS, THERE IS NO RECONSTRUCTION INTO A-POSITIONS), ALTHOUGH LOTS OF EVIDENCE EXIST FOR A'-MOVEMENT.

Consider (32):

(32)

(i) a proof is likely [_{<_>} <t> to be [discovered <t>]]

(ii) who did you say [<t> has [<t>' discovered the proof]]

In (i), each <t> = <a proof>, and the chain headed by the matrix subject is {P<1>, P<2>, P<3>}, the three occurrences of <a proof> (P<1> = "is likely _," etc.). In (ii), on the assumptions of <MP>, the external system at the LF interface requires an operator-argument analysis, with <who> "reconstructed" to an operator binding the (also "reconstructed") A-chain argument headed by <t>.

WE WILL TALK ABOUT THIS LATER ON IN THE SEMESTER, AND NOTE ALSO:

{See <MP>, chap. 3.5, and for improved versions with many consequences, Fox (1998), Sauerland (1998b). A question that might be raised is where the "reconstruction" operation takes place: within narrow syntax or on the other side of the interface, along with Binding Theory (I am assuming, following <MP>) and other interpretive systems.

WE WILL TRY TO DEAL WITH RECONSTRUCTION AND BINDING IN THESE TERMS, AS POST-LF PROCESSES.

Similar questions arise about other covert operations; see note <TYP>.><RECON = footnote>

THIS IS A REFERENCE TO FOOTNOTE 45, ON STYLISTIC PROCESSES AND OTHER POST-CYCLE STUFF. IN A SENSE, THIS PAPER IS TAKING THE CYCLE AS A CRITERION FOR NARROW SYNTAX.

A straightforward implementation can be based on analysis into two chains: the A'-chain (<who>, <t>), and the A-chain (<t>, <t>').

WELL, LET'S SEE HOW STRAIGHTFORWARD THE IMPLEMENTATION IS. TO BEGIN WITH, WE MUST WRESTLE WITH CONSERVATION MATTERS. HOW CAN WE GET TWO CHAINS OUT OF A SINGLE OCCURRENCE OF who?

TWO POSSIBILITIES. FIRST, who IS MORE COMPLEX THAN IT LOOKS. GOING BACK

TO IDEAS BY KLIMA AND KURODA, WE MAY THINK OF THIS AND SIMILAR ELEMENTS AS INVOLVING, IN EFFECT, TWO FORMATIVES: AN OPERATOR AND A PREDICATE/VARIABLE. THIS IS MORE APPARENT IN *which one*, AND SIMILAR ELEMENTS. IF THAT'S THE WAY TO GO, ALTHOUGH OVERTLY WE HAVE THE SINGLE LEXICAL ITEM *who*, IN SOME SENSE (PERHAPS THROUGH LEXICAL FEATURES), WE NONETHELESS HAVE ENOUGH MATERIAL TO COME UP WITH TWO CHAINS, SINCE IN EFFECT WE'RE NOT GAINING INFORMATION, BUT SIMPLY SHIFTING IT AROUND. FOR CONCRETENESS, LET'S SAY THAT IN THE NUMERATION WE HAVE (THE OPTION OF) TWO FEATURES THAT NEED NOT BE SEEN AS PERTAINING TO A SINGLE LEXICAL ITEM, AND EVENTUALLY LF RECOVERS EACH FEATURE AS CORRESPONDING TO A SEPARATE CHAIN. NOTE, INCIDENTALLY, THAT THE CHAINS CANNOT BE TOTALLY DISTINCT, FOR AFTER ALL ONE STARTS WHERE THE OTHER ONE ENDS. WHETHER THAT'S DEEP DEPENDS ON WHAT ONE DOES WITH RESUMPTIVE PRONOUNS, WHERE CLEARLY THE TWO CHAINS ARE LEXICALLY SEPARATE (i.e. *Who do you wonder whether Mary likes him?*). THERE, TOO, YOU WANT THE TWO CHAINS TO BE CONNECTED, BUT THIS TIME THE DERIVATION IS NOT GOING TO HELP YOU, SO PERHAPS YOU NEED SOMETHING LIKE HIGGINBOTHAM'S ANTECEDENCE.

A SECOND POSSIBILITY IS A BIT MORE RADICAL. IT COULD BE THAT EXTRA CHAINS CAN EMERGE IN THE COURSE OF THE DERIVATION UNDER CERTAIN (HOPEFULLY LIMITED) CONDITIONS. IN THAT CASE YOU WOULDN'T TRY TO GET TWO CHAINS OUT OF TWO SEPARATE NUMERATION ITEMS, BUT RATHER YOU'D ADMIT THE POSSIBILITY THAT N ITEMS IN THE NUMERATION PRODUCE N+M CHAINS. WE WILL RETURN TO THIS ISSUE WHEN WE ANALYZE NUNES'S TREATMENT OF PARASITIC GAPS, PARTICULARLY IN THE TERMS DISCUSSED IN NUNES & URIAGEREKA (1999)--HENCEFORTH N&U.

THAT'S WITH REGARDS TO CONSERVATION, BUT THERE'S MORE ABOUT CHOMSKY'S IMPLEMENTATION:

The three occurrences of *<who>* are analyzed into the chains {Q<1>, Q<2>} and {R<1>, R<2>}, where Q<1> = "did you say....the proof," Q<2> = R<1> = "has...the proof," and R<2> = "discovered the proof." This is one of the permissible analyses into chains; it need only be permitted, not forced. In (i) there is a three-membered chain, in the (interpretable) analysis of (ii), there are two two-membered chains, sharing one occurrence of *<who>*.

THIS IS WHAT WE MUST STUDY--WHY IN SOME INSTANCES YOU GET A SINGLE CHAIN (A-MOVEMENT, PERHAPS SUCCESSIVE CYCLIC, THOUGH NOT OBVIOUS) WHEREAS IN OTHERS YOU GET AT LEAST TWO (PERHAPS AT MOST TWO) CHAINS (A'-MOVEMENT AND A-MOVEMENT). AT ISSUE, IN PART, IS WHETHER INTERMEDIATE RECONSTRUCTION SITES EXIST (FAIRLY OBVIOUS FOR A'-MOVEMENT, LESS CLEAR FOR A-MOVEMENT).

{Attempts in *<MP>* to account for successive-cyclic movement in terms of linked chains and other devices can be eliminated; they were based on failure to take seriously enough the actual

notion of <chain>.

I DON'T THINK I UNDERSTAND THIS, AND THE COMMENT IS TOO CRYPTIC TO ASCERTAIN. CLEARLY THE MP SYSTEM WAS RATHER CUMBERSOME, BUT IT IS NOT OBVIOUS TO ME WHAT IN THE PRECISE NOTION OF CHAIN MAKES IT SO EASY TO COLLAPSE THE CHAIN HERE OR THERE. AGAIN, WE WILL RETURN TO THIS ISSUE LATER ON.

They were also in error, in not recognizing the role of intermediate traces in interpretation and computation.

THAT'S CERTAINLY TRUE, AT LEAST FOR A'-MOVEMENT. NONETHELESS, WE'LL HAVE TO SEE WHETHER THIS ERROR IS AMENDED IN THE PRESENT SYSTEM.

The conclusions drawn there about trace invisibility no longer hold, though interesting aspects of the question remain. }

AGAIN, TOO CRYPTIC A COMMENT TO INTERPRET--SO LET'S POSTPONE IT.

A chain, then, is a set of occurrences of an object $_$ in a constructed syntactic object K . It would make sense to rethink all notions involving chains in these terms, including their interpretation at the interface. If we do, some apparent problems disappear. Suppose that raising of DP to SPEC-T checks and deletes its uninterpretable Case feature. We want DP and its trace $\langle t \rangle$ to be identical, so the feature must also delete in the trace. But what guarantees that the feature is deleted throughout the chain? The question does not arise if we think of a chain, more properly, as a set of occurrences of $_$ in K : the feature is deleted in the single element $_$, unproblematically.

AN ANALOGY MIGHT HELP. SUPPOSE YOU'RE TELEPORTED FROM YOUR HOME TO THE ANDROMEDA GALAXY, AND UPON ARRIVAL SOME CREATURE SNAPS OFF YOUR MEMORY FOR YOUR GRANDMOTHER'S FACE. SAY YOU GO ABOUT YOUR BUSINESS WITHOUT KNOWING THIS, AND THEN YOU'RE TELEPORTED BACK. THEN YOU SEE GRANNY AT HOME; DO YOU REMEMBER HER FACE? WELL, IF THE TRICK IS PLAYED THE WAY I SUGGESTED BEFORE, THE ANSWER WOULD BE NO, EVEN IF IT WASN'T AT HOME THAT THE MEMORY GOT SNAPPED. (INCIDENTALLY, YOU COULD HAVE USED GORIER SCENARIOS, LIKE SNAPPING AN EAR, SAY.) ANYWAY, I DON'T THINK ANY OF THIS IS AS 'UNPROBLEMATIC' AS CHOMKSY SUGGESTS, FOR IT FORCES YOU INTO A FASCINATING ONTOLOGY. (NOT THAT I HAVE PROBLEMS WITH THAT ONTOLOGY, BUT A RATHER SERIOUS SUSPENSION OF DISBELIEF IS NECESSARY HERE.)

Should the notions of occurrence and chain be extended to proper subparts of LI that are not elements of the lexical array: features and sets of features? Let's put the question aside for the moment,

YOU WILL SEE THAT THIS ONE IS IMPORTANT IN THIS PAPER.

and turn to the relations permissible under Condition (Dii) of (30). Consider the operation Merge (indispensable in some form).

MIND YOU, FOR FACTUAL REASONS--NOT FOR OBVIOUS REASONS HAVING TO DO WITH THE STRONG THESIS. NEED FOR MERGE PRESUPPOSES EITHER THAT INTERPRETATION IS NECESSARILY FREGEAN (SAY) OR ELSE THAT A COMPUTATIONAL SYSTEM WHICH IS MORE COMPLEX THAN A MARKOVIAN SYSTEM IS FOR SOME REASON NECESSARY WHEN ONE INVESTIGATES THIS SORT OF STRUCTURE. IN BOTANY A SIMILAR QUESTION ARISES, INCIDENTALY, SINCE ALL VEGETABLE GROWTH PRESUPPOSES FRACTAL STRUCTURES, WHICH ARE NOT GENERABLE BY STANDARD MARKOVIAN SYSTEMS. PERHAPS WHATEVER UNDERLIES THAT (WHICH IS EXTENSIBLE, IT SEEMS, TO THE INNER PLANS OF ALL MULTICELULAR ORGANISMS) MIGHT BE RELATED TO WHATEVER IS NECESSARY FOR MERGE--AT LEAST IT IS WORTH ASKING WHAT THAT WOULD MEAN.

Merge takes two objects $_$ and β and forms a new object $K(_,\beta)$. The operation provides two relations directly: <sisterhood>, which holds of $(_,\beta)$, and <immediately contain>, which holds of $(K,_)$, (K,β) , and (K,K) (taking it to be reflexive).

THIS IS ALL TRUE, BUT NOT GIVEN. THOSE DEFINITIONS HAVE THE STATUS OF PRIMITIVES IN THE THEORY, HENCE WE MUST ASK THE USUAL QUESTIONS ABOUT THEM (WHY THOSE, ETC.).

Suppose we permit ourselves the elementary operation of composition of relations.

DITTO--NOTHING TRIVIAL, PERHAPS RELATED TO THE ISSUES MENTIONED BEFORE.

Applying it in all possible ways, we derive three new relations: (i) the transitive closure <contain> of <immediately contain>; (ii) identity (= (sister(sister))),

THIS IS IDENTITY ONLY ASSUMING THE BINARITY OF MERGE, WHICH IS ALSO MERELY ASSUMED--AND THIS ONE IS CERTAINLY NOT TRUE OF, SAY, FRACTAL SYSTEMS. KAYNE HAS A WAY OF DEDUCING BINARITY, BUT CHOMSKY SEEMS SKEPTICAL ABOUT KAYNE'S ARGUMENT ANYWAY, SO HE HAS TO STICK TO A COMPLETE STIPULATION. WITHOUT BINARITY, A SISTER OF A SISTER NEED NOT BE ITSELF (NOT THAT MUCH RELIES ON THIS, THOUGH).

and (iii) c-command (= sister(contain)).

THERE ARE A COUPLE OF TRICKS HERE. FIRST, THAT IS NOT WHAT WE GET BY APPLYING THE COMPOSITION OF ELEMENTARY RELATIONS. RATHER, WHAT WE GET IS (sister (immediately-contain)), WHICH DOESN'T HAVE THE LONG-DISTANCE CHARACTER OF COMMAND. SECOND, BY THE SAME REASONING WE GET (contain (sister)), WHICH DEFINES THE RELATION BETWEEN, SAY, A MAXIMAL PROJECTION

AND THE SISTER OF ITS HEAD. PERHAPS THAT'S THE WAY TO DEFINE THE ELEMENTS IN THE MINIMAL DOMAIN OF A GIVEN HEAD, BUT CHOMSKY DOESN'T EXPLORE THIS. I SHOULD SAY THAT, IN PART, ALL OF THIS IS AN ATTEMPT AT GETTING COMMAND AS A NATURAL RELATION WITHOUT INVOKING EPSTEIN'S DEDUCTION BASED ON KAYNE'S LCA. AS YOU'LL SEE LATER ON, THE LCA DOESN'T SQUARE WELL WITH SOME OF CHOMSKY'S EMPIRICAL ASSUMPTIONS, HENCE THE NEED TO DO THINGS WITHOUT IT. WE WILL RETURN TO CHOMSKY'S SKEPTICISM, TO SEE WHETHER WE CAN ADDRESS IT. PRESENTLY, I MUST SAY WE HAVE STIPULATED COMMAND, NOT THAT WE HAVE DERIVED IT FROM ANYTHING SERIOUS, AS IN EPSTEIN'S DEDUCTION (WHICH IS NOT WITHOUT ITS TROUBLES, BUT IS GENERALLY WELL TAKEN).

Thus K contains α if K immediately contains α , or immediately contains L which contains α ; conversely, α is a <term> of K if K contains

α .
THIS POSES SOME PROBLEMS FOR ADJUNCTION, BUT I WILL TRY TO STAY AWAY FROM THEM. IF YOU NEED A MORE FORMAL PRESENTATION OF THESE SORTS OF ISSUES, SEE THE APPENDIX TO RHYME AND REASON BY NUNES AND THOMPSON.

And α c-commands β if α is the sister of K, which contains β .

{The compositional definition of c-command is suggested by Epstein's derivational approach (see note <EPS>): α c-commands β if α is merged with K containing β . There are empirical differences between the two approaches, to which we return.

THAT'S WHAT I MENTIONED--THEY REVOLVE AROUND THE LCA.

The derivational definition also raises some questions: in particular, why does "containment" enter? (i.e., why does X merged with Y c-command terms of Y?).

THIS IS THE DIFFICULTY I HAD IN MIND WHEN MENTIONING EPSTEIN'S SYSTEM. SEE DRURY (1999) FOR A FULLER DISCUSSION.

The matter is addressed in Epstein et al. (1998, ch. 6), but inconclusively, as far as I can see. An argument for asymmetry of c-command is also presented there; I am assuming that there is no asymmetry, its effects being derived in other ways.}<DER = footnote>

THAT IS TO SAY, THERE IS NOTHING PRIVILEGED, IN CHOMSKY'S SYSTEM, ABOUT THE C-COMMANDING ELEMENT.

The relation of c-command is available, and expected, on very weak assumptions.

I MUST SAY THIS IS NOT OBVIOUS TO ME.

The relation has played a large role in syntactic theory, though it may be that it does not function within narrow syntax but only in interpretation of the information it provides -- that is, in

mapping it to syntactic objects that belong to mental systems external to the language faculty itself (see note <TYP>). That might be expected if external systems access representations (PF and LF) to which the notions <sister> and <contain> apply.

I WOULD FIND IT SURPRISING IF EXTERNAL SYSTEMS *BOTH* USE COMMAND (AS THEY APPEAR TO), PARTICULARLY SINCE THE ASSUMPTIONS WE'VE GIVEN FOR COMMAND ARE NOT WEAK, REALLY. THAT COMPOSITION WITH THE TRANSITIVE CLOSURE OF ONE OF THE BASIC RELATIONS THAT MERGE ALLOWS IN TERMS OF THE OTHER BASIC RELATION SHOULD HAPPEN TO BE THE WAY THAT DIFFERENT INTERFACE SYSTEMS WORK IS, TO SAY THE LEAST, A REMARKABLE EMPIRICAL FINDING. EPSTEIN'S APPROACH WAS CONCEPTUALLY MORE PLEASING: THE DERIVATION ITSELF GIVES YOU COMMAND, HENCE YOU EXPECT TO SEE IT IN BOTH SIDES OF THE GRAMMAR. OF COURSE, IT'S NOT COMPLETELY CLEAR THAT THE DERIVATION *SHOULD* GIVE YOU COMMAND, SINCE OTHER COMPUTATIONAL SYSTEMS WITH ROUGHLY THE SAME POWER INVOLVED IN MERGE CLEARLY DO NOT GIVE YOU COMMAND (E.G. YOU DON'T HAVE IT BOTANY OR GENETICS). IT IS A DEEPLY SURPRISING PROPERTY OF THE MERGE SYSTEM THAT GUYS THAT MERGE SHOULD CARE ABOUT THE DESCENDENTS OF THEIR SISTERS.

FOR WHAT IT'S WORTH, HERE'S A DIFFERENT STAB AT COMMAND. WHEN YOU MERGE THINGS, THE DERIVATION ITSELF DICTATES ONE OF TWO MODES: EITHER YOU ARE WITHIN THE SAME DERIVATIONAL WORKSPACE, OR ELSE YOU NEED TO SEPARATE WORKSPACES. FOR INSTANCE, IN he saw a woman YOU ONLY NEED ONE DERIVATIONAL WORKSPACE, WHEREAS IN [the man] saw a woman YOU CLEARLY NEED TWO: ONE TO ASSEMBLE the man AND ONE TO ASSEMBLE saw a woman, TO WHICH YOU MERGE the man IN DUE TIME. IT IS A SIMPLE FACT ABOUT DERIVATIONS THAT ELEMENTS THAT ARE ACTIVE WITHIN THE SAME DERIVATIONAL WORKSPACE STAND IN A COMMAND RELATION. IF THIS WERE THE WAY TO GO, WE WOULDN'T BE MAKING COMMITMENTS TO ANYTHING HAVING TO DO WITH LCA (ALTHOUGH THERE SURELY WOULD BE CONSEQUENCES FOR THE LCA IF THIS WERE TO BE ASSUMED), AND IT WOULDN'T BE SUCH A BIG SURPRISE THAT PF AND LF SHOULD CARE ABOUT COMMAND UNITS, SINCE THESE WOULD BE CARVED UP BY THE DERIVATION ITSELF. IN ESSENCE, THAT'S THE MSO SYSTEM IN THE NARROW SENSE.

The sisterhood relation is significant primarily (perhaps only) for heads, that is, LIs and M(odified)LIs formed from them. Furthermore, sisterhood relations presumably remain if LI is modified to MLI: if VP is sister of T, for example, it should remain so even if uninterpretable features are deleted from T. More generally, LI and its modifications are not distinguished with regard to the fundamental relations defined in terms of Merge.

BASICALLY THIS TREATS ADJUNCTION TO HEADS AS HAPPENING ON A DIFFERENT PLANE OF ITS OWN, SINCE IT DOESN'T AFFECT ANYTHING IN THE PHRASE-MARKER. (UNLESS CHOMSKY ISN'T TALKING ABOUT HEAD ADJUNCTION AT ALL, GIVEN HIS NEW NOTION OF 'AGREE'; THAT IS, IN MP, THE WAY TO CHECK

A FEATURE IN A HEAD H WAS TO ADJOIN ITS CHECKER TO H, BUT NOW WE WILL DO THE TRICK LONG-DISTANCE, DIRECTLY ELIMINATING THE RELEVANT FEATURE WITHIN H; IF SO, MLI MIGHT JUST BE THE MODIFIED HEAD, WITH NO ADJUNCTS--THEN AGAIN ONE HAS TO WORRY ABOUT SPELLING OUT THE DETAILS OF WHAT LOOKED LIKE GENUINE ADJUNCTION, E.G. ON HEAD MOVEMENT, WHICH WE RETURN TO BELOW.)

For $_$ an LI, then, we extend the relations defined for $(_,\beta)$ to $(MLI(_),\beta)$, MLI constructed from $_$ (typically -- perhaps only -- by deletion of uninterpretable features). We extend the same convention to a feature $F(_)$ of a head $_$. Thus if LI $_$ is the sister of β or c-commands β , then $MLI(_)$ and $F(_)$ do as well.

I AM NOT HAPPY ABOUT SIMPLY 'EXTENDING CONVENTIONS', IF THESE WILL HAVE EMPIRICAL CONSEQUENCE--THE MINIMALIST MUST JUSTIFY THE EXTENSION. IN ANY CASE, IT SEEMS AS IF THIS IS REALLY SAYING THAT CONFIGURATIONAL RELATIONS ESTABLISHED UPON MERGE ARE KEPT EVEN AFTER THE TRANSFORMATIONAL COMPONENT OPERATES. IF SO, THIS IS INTERESTING, FOR THE OBVIOUS QUESTION IS WHY SHOULD MERGE RELATIONS PERSIST THAT WAY? COULD IT BE THAT THIS IS ANOTHER CONSERVATION? IF SO, NOT ONLY WOULD WE HAVE THE LEXICAL CONSERVATION LAW THAT WE SAW BEFORE, BUT ALSO A PHRASAL CONSERVATION LAW. IN TURN, THIS WOULD ENTAIL THAT PHRASES DEFINE THEIR OWN SEPARATE NATURAL CLASS, JUST AS WORDS DO. EMPIRICALLY THIS SEEMS SOUND, AND IT SHOULD BE POINTED OUT THAT IT HAS THE EFFECT OF CREATING A WORD-LEVEL COMPONENT (WHERE LEXICAL CONSERVATION HOLDS AS AN INPUT-OUTPUT CONDITION ON DERIVATIONS) AND A PHRASAL-LEVEL COMPONENT (WHERE PHRASAL CONSERVATION AGAIN HOLDS AS A FURTHER INPUT-OUTPUT CONDITION ON DERIVATIONS). THE LATTER, OF COURSE, IS VERY REMINISCENT OF THE TRADITIONAL MAPPING BETWEEN D-STRUCTURE AND LF, BUT THAT ONE WAS DONE IN TERMS OF LEVELS OF REPRESENTATION (UNIFIED OBJECTS), WHEREAS THE PRESENT CONDITION IS ONE THAT HOLDS OF CHUNKS OF DERIVATIONS, NOT UNIFIED LEVELS. I FIND THIS IMPORTANT IN ESTABLISHING THE CLAIM THAT UNDERLIES MUCH OF MY COMMENTARY TO THIS PAPER: THAT WHEN YOU LOOK CAREFULLY, DISECTING THE ASSUMPTIONS, CHOMSKY IS MAKING THE SUBSTANTIVE CLAIM THAT SOMETHING LIKE A D-STRUCTURE COMPONENT (CLEARLY, NOT A LEVEL) HOLDS IN THE SYSTEM. AT THIS POINT, THE ISSUE IS PRESERVING THE PHRASAL INTEGRITY OF THE SYSTEM. IT SHOULD BE CLEAR, I HOPE, THAT A TRANSFORMATIONAL SYSTEM COULD OCCUR WHICH LITERALLY DESTROYS PHRASAL INTEGRITY. INDEED, NATURAL SYSTEMS OUT THERE PROBABLY EXIST WITH THOSE PROPERTIES, WITH RADICAL TRANSFORMATIONS DESTROYING, IN THE PROCESS OF APPLYING, WHAT WENT INTO THEM. BUT LINGUISTIC TRANSFORMATIONS, AT LEAST CORE ONES, ARE STRUCTURE PRESERVING--PERHAPS THE MAJOR RESULT OF JOE EMMONDS'S WORK.

Questions have arisen about the interpretation of these notions for adjunction, particularly

head-adjunction of H' to H.

THAT'S WHAT I WAS MENTIONING ABOVE--SO CHOMSKY DOES AT LEAST ACKNOWLEDGE THE ISSUE.

With no further elaboration, H and H' would be sisters and neither would c-command outside.

THAT'S THE CASE IN THE TRADITIONAL SYSTEM.

Whether this matters is unclear. Failure of H' to c-command its trace seems to have no significance in the present framework;

THAT'S THE CASE, IN LARGE PART, BECAUSE THE PRESENT FRAMEWORK HAS BASICALLY NOTHING TO SAY ABOUT INSTANCES WHERE C-COMMANDING A TRACE SEEMED TO BE IMPORTANT. WE WILL RETURN TO SOME OF THESE.

failure of H to c-command into its former c-command domain would be problematic only if H functions after adjunction in implementing agreement and movement, but that does not seem necessary.

THAT'S AN INTERESTING QUESTION. TAKE INFL AFTER A VERB HAS MOVED TO IT. DOES INFL (NOT INFL+VERB) FUNCTION IN ANY RELATION OF GRAMMAR? PERHAPS CLITICIZATION IN ROMANCE IS OF THIS SORT, SINCE CLITICS SEEM TO GO TO AN INFL-LIKE ELEMENT REGARDLESS OF WHETHER V HAS MOVED. THE CASE IS TOO MURKY TO GO INTO NOW, BUT I'LL RETURN TO EXAMPLES OF THIS SORT.

Pending some good reason to sharpen the sisterhood relation for this case, I will leave it as is.

{Sisterhood relations would be carried over under head-adjunction if the result were taken to be an extended MLI.

WHICH IS THE OTHER ISSUE I WAS RAISING ABOVE. PERHAPS SOME INSTANCES ARE OF THIS SORT (WHAT SOME OF US CALLED 'MORPHOLOGICAL INCORPORATIONS') WHEREAS OTHERS ARE NOT ('FUNCTIONAL INCORPORATIONS')

Many questions dissolve if head-adjunction is part of the phonological component. There are, I think, independent reasons to suspect that this may be true, at least over an interesting range, but I will defer the question.}<HADJ = footnote>

THIS IS WHAT I MENTIONED ABOUT V-2, ALTHOUGH IT IS MUCH LESS CLEAR FOR, SAY, V TO I.

Though varieties of government would be "imperfections," to be avoided if possible, the closer-to-primitive notion of L-marking should pass muster, hence also notions of barrier that are

based on nothing more than L-marking.

WELL, THIS IS A BIT FAST, PARTICULARLY IF YOU RECALL THE DEFINITIONS OF L-MARKING OF TEN YEARS AGO. R&U HAD ONE DEFINITION THAT MIGHT INDEED (WITH SOME MODIFICATIONS TO UPDATE IT) HAVE THE RELEVANT EFFECTS, BUT I WON'T GO INTO THIS NOW.

Here numerous questions arise about island conditions within a minimalist framework, about which I have nothing useful to say. {See Kitihara (1998), Lasnik (1995a,b), Uriagereka (forthcoming a).} <URIA = footnote>

WE WILL TRY TO RETURN TO THIS, SINCE THE HIGHLY DERIVATIONAL SYSTEM IS A RATHER NATURAL CANDIDATE FOR EXPLAINING EXTRACTION DOMAINS.

<5. Imperfections>

There are some respects in which the strong thesis seems untenable, and we find what appear to be "design flaws" that are not necessary for language-like systems. The most obvious involve the phonological component, which takes syntactic objects constructed by the computational operations C<HL> and converts them to representations at the PF interface. Here there are radical violations of the interpretability and inclusiveness conditions ((B) and (C) of (30)). Condition (C) is violated by operations that introduce such new elements as prosodic structure and narrow phonetics. Condition (B) is violated by the discrepancy between the phonological properties of lexical items LI ("morphophonemes," "phonological units," etc., within various frameworks and terminologies) and the narrow phonetic instantiations of combinations of such elements. It may be that phonological features of LIs do not even appear at the level PF, that the "input" and "output" of the phonological component are in different "vocabularies." In that case the interpretability condition would be maximally violated by the operations of the phonological component, while the inclusiveness condition is clearly inoperative.

IN OTHER WORDS, THIS COMPONENT WOULD NOT BE SUBJECT TO A STRONG CONSERVATION LAW OR THE GENERAL DESIGN INTUITION THAT FL HAS PROPERTIES THAT ARISE ONLY AT *MERE* INTERFACES--ALTERNATIVELY:

We therefore turn to the alternatives of (29). Option (ii) seems implausible; the properties appear to be real, whatever their proper expression. Of the two remaining possibilities, the most interesting by far is (iii): the properties are real, but not imperfections: the super-engineer, called upon to map independently-motivated syntactic objects to PF, would hit upon the phonological component as an optimal solution. I have no idea whether this can be formulated as a sensible research task.

DITTO...

As for option (i), it might be argued that departures from "good design" are not surprising in this

domain. Direct evidence about sound systems is available for language acquisition, permitting a degree of complexity and variation. And the subsystem reflects special properties of the sensorimotor systems, which are in a certain sense "extraneous" to language, relating to externalization by systems with non-linguistic properties, and capable of much variation while FL remains fixed, as in sign languages. Symbolic systems designed for special purposes (metamathematics, computers, etc.) dispense entirely with a phonological component, not facing the need to meet the legibility conditions for human language at the sensorimotor interface.

IT IS WORTH POINTING OUT, PERHAPS, THAT SYMBOLIC SYSTEMS ARE TYPICALLY EXPRESSED IN WRITING, WHICH IMPOSES SOME LIMITATIONS WHICH ARE COMPLETELY EXTRANEIOUS TO THE SYSTEM. FOR INSTANCE, YOU NEED SUFFICIENTLY DISTINCT SYMBOLIC REPRESENTATIONS, VISUALLY SPEAKING, SO THAT YOU DON'T CONFUSE "PLUS" OR "MINUS" AND SO FORTH, AND THESE ARE COMPLETELY ARBITRARY AND PRESENT OBVIOUS VARIATION ACROSS CODING SYSTEMS. THERE IS A SENSE IN WHICH THOSE PROPERTIES OF THE SYSTEM ARE COMPLETELY IRRELEVANT TO ITS INTERNAL WORKINGS, AND MUCH AS I HATE TO SAY THIS CLEARLY, IT COULD WELL BE THAT SO TOO ARE PHONETIC AND EVEN SOME PHONOLOGICAL PROPERTIES OF THE LANGUAGE FACULTY; I DON'T MEAN THIS IN A DIMINING WAY, THOUGH, IN TWO SENSES. FIRST OF ALL, THE SYSTEM IS WHATEVER IT IS, AND IF IT HAS PROPERTIES THAT ARE OUTSIDE WHAT WE CALL "FL" IN A TECHNICAL SENSE, WELL--SO BE IT. BUT MOST IMPORTANTLY FROM A SYNTACTIC POINT OF VIEW, THERE PROBABLY ARE SOME PROPERTIES OF THE PHONOLOGICAL SYSTEMS THAT AFFECT SYNTACTIC COMPUTATION. FOR INSTANCE, AS WE WILL SEE AT LENGTH LATER ON IN THE COURSE, LINEARIZATION REQUIREMENTS (HIGGINBOTHAM'S 1983 POINT) AND LIMITATIONS IMPOSED BY THE MORPHOLOGICAL SYSTEM (IN MARANTZ'S OR EVEN MORE RADICALLY IN ANDERSON'S SENSE). TO GO ON WITH THE ANALOGY WITH SYMBOLIC SYSTEMS, SUCH LIMITATIONS EXIST THERE AS WELL, FOR INSTANCE THE CONVENTION IN FORMAL LANGUAGES THAT ELEMENTS THAT ENTER INTO A RELATION ARE WRITTEN CONTIGUOUSLY, OR WITH A MARKER FOR THE RELATION IN BETWEEN THEM--NOT LONG DISTANCE. IT'S NOT AS IF, MATHEMATICALLY, YOU COULDN'T EXPRESS A MORE LONG-DISTANCE RELATION (FOR INSTANCE IN A CONVOLUTED GRAPH, OR A QUANTUM ONE--AND GET READY FOR THAT IN QUANTUM COMPUTATION!) BUT NO FORMAL SYSTEM WORKS LIKE THAT (I DON'T CLAIM TO KNOW WHY, ALTHOUGH THIS IS PROBABLY A REFLEX OF THE WRITTEN FORMAT, PERHAPS PIGGY-BACKING ON VISUAL AND EVEN PSEUDO-LINGUISTIC DEMANDS). MY POINT IS: SOME ASPECTS OF THE "PUBLIC SUPPORT" (IN THE SENSE OF WHAT MAKES THE SYSTEM OF KNOWLEDGE COMMUNICABLE) COULD WELL LIMIT THE FORMAT OF THE SYSTEM. NOW THAT'S INTERESTING. CURIOUSLY, THOUGH, IT IS MUCH LESS CLEAR, WITH REGARDS TO THOSE, SHALL WE CALL THEM, "DEEP" PROPERTIES OF THE PHONOLOGICAL SYSTEM THAT THEY DO VIOLATE INCLUSIVENESS AND INTERPRETABILITY. IN ESSENCE, THAT'S THE QUESTION OF "FAITHFULNESS" IN OT, ON THE ONE HAND, AND AGAIN IN OT, OF THE NATURE OF PHONOLOGICAL CONSTRAINTS. ABOUT THE LATTER, LITTLE IS KNOWN, SO IT'S HARD TO SPECULATE. BUT THE ISSUE WOULD BE WHETHER ANY OF THE

"DEEP" (IN THE SENSE OF FL-RELATED) CONDITIONS OF THE PHONOLOGICAL COMPONENT ARE OR ARE NOT THERE FOR INTERPRETIVE REASONS. IN ANY CASE, ANSWERING THAT DOESN'T SEEM MORE OR LESS CLEAR THAN ANSWERING A SIMILAR QUESTION, SAY ABOUT CHAINS, IN THE LF SIDE.

FOR INSTANCE, STRESS ASSIGNMENT, WHICH WE KNOW IS ALTERED BY SYNTACTIC CONSIDERATIONS. IS THAT PHENOMENON THERE FOR INTERPRETIVE REASONS? PERHAPS, IT DEPENDS ON WHETHER YOU BELIEVE THE OLD IDEA THAT STRESS AIDS OR EVEN DETERMINES WORD IDENTIFICATION. SUPPOSE IT DOES; THEN YOU CAN THINK OF IT AS A WAY TO MATCH PF OUTPUT TO THE LEXICAL ARRAY YOU START WITH. THAT, NOTE, WOULD INDICATE THAT SUCH A MAPPING SATISFIES SOME FORM OF INCLUSIVENESS, THE OTHER ISSUE THAT CONCERNS US. AND HOW ABOUT STRESS ALTERATIONS, IMPOSED BY RHYTHMIC CONSIDERATIONS? AGAIN, IT DEPENDS ON WHAT RHYTHM IS DOING IN THE SYSTEM, AND HOW MUCH THOSE ALTERATIONS AFFECT *PRIMARY* STRESS, WHICH SEEMS MUCH MORE CONSERVATIVE THAN SECONDARY (AND TERTIARY, ETC.) STRESS. PERHAPS RHYTHM IS A PURE INTERPRETIVE IMPOSITION OF THE PF COMPONENT, RELATED TO BREATHING, DISCERNABILITY, AND/OR WHAT NOT (I DON'T THINK THESE QUESTIONS HAVE BEEN ASKED IN THIS WAY, SO I DON'T REALLY KNOW). IF SO, IT MAY PERHAPS BE LEGITIMATE FOR RHYTHM TO ALTER STRESS ASSIGNMENT, ALTHOUGH THEN PERHAPS ONLY PRIMARY STRESS OBEYS THE CONSERVATION LAW (AND INCIDENTALLY, ONLY PRIMARY STRESS WOULD SEEM TO RELATE TO TOKENS IN THE NUMERATION; SECONDARY ONE IS CLEARLY A PURELY RHYTHMIC DEVICE, SO PERHAPS THE ISSUE OF CONSERVING IT SIMPLY DOESN'T ARISE). MY POINT: DEEP PHONOLOGICAL PROPERTIES MIGHT WELL BE MUCH MORE ELEGANT THAN IT SEEMS.

OTHER PHONOLOGICAL PROPERTIES ARE PERHAPS "INELEGANT", IN THE TECHNICAL SENSE, BUT SO WHAT? "INELEGANCE" IS WELCOME IN AN UNDERSPECIFIED SYSTEM--INDEED, IT IS OFTEN WHAT MAKES IT VIABLE IN OUR OTHERWISE TOO REGULAR UNIVERSE. WITHOUT BUMPS IN THE EXPANSION OF THE UNIVERSE EVERYTHING WOULD BE HARMONICALLY NOT HERE!

The strongest sustainable inclusiveness/interpretability requirement, then, is (33):

(33) Inclusiveness holds of narrow syntax, and each feature is interpreted at the level LF or associated with phonetic features by the phonological component

PERSONALLY, I FIND THIS A BIT TOO WEAK. OF COURSE, SINCE CHOMSKY WILL BE GIVING UP THE LCA (WHICH FALLS INTO WHAT I WAS SAYING ABOVE), AND DOESN'T SEEM TO CARE TOO MUCH ABOUT MORPHOLOGICAL DETAIL, (33) IS EXACTLY WHAT HE WANTS. BUT WE'LL BE SEEING, FOR INSTANCE, CLITICIZATION EFFECTS OF THE SORT RAPOSO HAS BEEN RECENTLY STUDYING, MUCH IN THE SPIRIT OF WAKERNAGEL REVISITED BY ANDERSON. THAT SYSTEM, FOR INSTANCE, SEEMS EXTREMELY ELEGANT AND WELL BEHAVED IN ITS CORE

ASPECTS, AND (33) WOULD RENDER IT IRRELEVANT.

The phonological component is generally assumed to be isolated in even stronger respects: there are <true> phonological features that are visible only to the phonological component and form a separate subsystem of FL, with its own special properties. Assume this to be true.

SURE, BUT THOSE PROPERTIES MIGHT BE WHAT I WAS CALLING SURFACE ONES. STRESS, INTONATION PHRASING, MORPHOLOGICAL UNIFICATIONS, AND OTHERS, DON'T SEEM TO BE SO ISOLATED.

Then in the course of construction of LF, an operation Spell-Out delivers the structure already formed to the phonological component, which converts it to PF.

RIGHT, WHICH IN AND OF ITSELF MEANS NOTHING, REALLY. IT IS ONLY A MATTER OF PERSPECTIVE TO TAKE SPELL-OUT AS SHIPPING STUFF OUT OF THE NARROW SYNTAX PATH. WE COULD ALSO THINK OF SPELL-OUT AS SHIPPING STUFF OUT OF THE SURFACE PATH, INTO THE LF BOX, OR (MORE PLAUSIBLY, GIVEN THE FACTS) AS SPELL-OUT LITERALLY *SPLITTING* A UNIFIED PATH IN TWO. I MEAN, THE FACT THAT THERE ARE AFTER-LF PROPERTIES, AND THUS ISOLATED ONES (SAY, CONDITIONS ON CONTEXT CONFINEMENT, TO BE CONCRETE) DOESN'T IN ITSELF ENTAIL THAT THE LF COMPONENT IS ISOLATED; MAYBE IT IS (WE THINK NOT), MAYBE IT ISN'T (WE THINK SO). ASSUMPTIONS HERE ARE OF LITTLE HELP, PARTICULARLY SINCE WE'RE MODIFYING THEM AS WE GO ALONG; THUS, FOR INSTANCE, MATTERS OF ANTECEDENCE WERE ONCE TAKEN TO BE PURELY SYNTACTIC, ALTHOUGH NOW THEY SEEM MORE POST-SYNTACTIC--WHY? BECAUSE ASSUMING THEY ARE SYNTACTIC MAKES OUR LIFE HARD FOR THE CONSERVATION AND INTERPRETABILITY ASSUMPTIONS. WITH THE SAME KIND OF REASONING WE COULD TECHNICALLY DECIDE THAT WHATEVER IS CONSERVATIVE AND OBEYS INTERPRETABILITY IN THE PF SIDE IS COMPLETELY WITHIN THE SYSTEM. IT'S EITHER THAT OR ELSE ABANDONING THE HOPE OF ACCOUNTING FOR, SAY, CLITIC PLACEMENT IN MINIMALIST TERMS.

If lexical items LI express Saussurean arbitrariness in the conventional way, then Spell-Out "strips away" the true phonological features, so that the derivation can converge at LF; it will crash if later operations introduce LIs with phonological features.

THAT'S ALL WELL, BUT EQUALLY LOADED. NOTE ONE MORE THING, INCIDENTALLY. IF YOU DON'T SAY SOMETHING ABOUT CONSERVATION IN THE PF SIDE, THEN IT'S WITTGENSTEINIAN BEDLAM ALLOVER AGAIN. THAT IS, NOTHING SHOULD PREVENT, IN PRINCIPLE, A SYSTEM THAT ALLOWS YOU TO START IN A NUMERATION WITH TEN ITEMS AND BY THE END OF THE DAY YOU PRONOUNCE THAT "ba". THE LINE I WAS SKETCHING ABOVE WAS VERY DIFFERENT. PRIMARY STRESS, FOR INSTANCE, IS IN THE SYSTEM TO ALLOW YOU TO IDENTIFY NUMERATION TOKENS (ASSUMING A SUBSTANTIVE NUMERATION). SIMILAR STORIES COULD BE TOLD ABOUT PHONOLOGICAL PHRASING, PERHAPS

A REFLEX OF NUMERATION-PHASES (AND AS SUCH AGAIN A WAY OF CODING THE INPUT-OUTPUT MAPPING BETWEEN THE NUMERATION AND THE PART OF THE SIGNAL YOU ACTUALLY GET AT PF). NEEDLESS TO SAY, WE KNOW THERE ARE LOTS OF READJUSTMENTS--INDEED, WE HOPE THAT THESE TO SOME EXTENT GEAR CLITICIZATION, IN ANDERSON'S SPIRIT. BUT THOSE READJUSTMENTS ARE PERFECTLY CONSERVATIVE ASWELL, HAPPENING AT THE EDGES AND THROUGH SIMILAR RECOVERABLE DEVICES. IN A WORD, IT SEEMS AS IF THE DEEP PHONOLOGICAL SYSTEM IS WELL DESIGNED TO ALLOW YOU TO GRASP NOT JUST NUMERATION TOKENS, BUT EVEN NUMERATION CHUNKS THAT ARE RELEVANT TO SYNTACTIC COMPUTATION (IN A SENSE TO BE DISCUSSED BELOW IN THE PAPER). AND HOW COULD IT BE OTHERWISE, FROM THE PERSPECTIVE OF ACQUISITION? CHOMSKY WILL INSIST ON VARIOUS PLACES ON THE IDEA THAT LINGUISTIC PROPERTIES ARE NOT IMPOSED BY LEARNABILITY (OR OTHER) CONSIDERATIONS. BUT IF THIS IS SO, THEN, THE PROPERTY I'M ALLUDING TO HAS TO BE A DESIGN PROPERTY OF THE MAPPING TO THE PF COMPONENT. AT LEAST, I DON'T SEE ANY OTHER ALTERNATIVE (OF COURSE, YOU COULD SHRUG YOUR SHOULDERS AND SAY THAT THE FACT THAT STRESS COINCIDES WITH NUMERATION TOKENS IS A MERE COINCIDENCE; BUT THEN AGAIN YOU COULD SHRUG YOUR SHOULDERS ABOUT LESS OBVIOUS THINGS: CHAINS, FOR INSTANCE).

On the assumptions of Distributed Morphology, the phonological features are introduced after Spell-Out by phonological operations applying to LIs lacking them.

THAT'S VERY PLAUSIBLE, AND INDEED PERHAPS THE WAY TO SEPARATE DEEP PROPERTIES OF PF FROM SHALLOW ONES. FOR INSTANCE, I'M PERFECTLY HAPPY WITH SAYING SOMETHING LIKE THIS: THE CONSERVATION OF NUMERATION-TO-PF INFORMATION DEMANDS INTONATION STRESS AND INTONATION PHRASING CORRESPONDING TO PHASES, BUT NEITHER OF THESE IS "ASSIGNED" TO WORDS. RATHER, YOU STRESS AN ABSTRACT WORD SHELL AND BRACKET AN ABSTRACT PHASE. LATER ON, DISTRIBUTED MORPHOLOGY DOES ITS REAL JOB, ASSIGNING ABSTRACT STRESS WHERE IT CORRESPONDS, READJUSTING FOR RHYTHM WITHIN A PHASE, AND SO ON (POSSIBLY MORE RADICALLY IN THE CASE OF EDGE-ORIENTED "STYLISTIC" DISPLACEMENTS).

I will assume some instantiation of this array of options to be correct.

Narrow syntax

BOTTOM LINE: THIS IS WHAT CHOMSKY IS INTERESTED IN HERE

also involves devices that are imperfections unless shown to be unreal (option (ii) of (29), which again seems implausible) or to be motivated by design specifications (option (iii) of (29), the most interesting possibility). Consider two striking examples:

- (I) Uninterpretable features of lexical items
- (II) The "dislocation" property

Under (I), we find features that receive no interpretation at LF and need receive none at PF, hence violating any reasonable version of the interpretability condition (B). The example that has played the most important role ever since Jean-Roger Vergnaud's famous unpublished letter twenty years ago is structural Case. The picture is more complex for agreement features: semantically interpretable for nouns, but not for verbs/adjectives, and phonetically optional throughout.

{There are important distinctions among these features, which I will largely ignore here. Problems of interpretation are also not trivial, as in:

- (i) animal languages (is, are) their main research interest
- (ii) three books (is, are) too much to read in a week
- (iii) we expected animal languages to be their main research interest
- (iv) we found three books too much to read in a week
- (v) animal languages raise(s) serious issues, seem(s) to be their main research interest)

In (i),(ii), the difference in meaning appears superficially to lie in the verbal inflection,

THAT IS, IN (i) FOR INSTANCE, THE PLURAL READING INVOKES REFERENCE TO VARIOUS LANGUAGES, WHEREAS THE SINGULAR ONE INVOKES REFERENCE TO A TOPIC.

but it carries over to forms lacking that inflection, as in (iii),(iv). And though agreement properties of the copula (with surface subject or complement) are a factor, the matter is more convoluted ((iv),(v)).

IT IS, ALTHOUGH A TREATMENT IS POSSIBLE WHERE THE COMPLEXITY OF THE SUBJECT IS WHAT TRIGGERS THE AGREEMENT WHEN AGREEMENT OBTAINS. WE WILL SEE A SYSTEM OF THIS SORT IN CASTILLO'S WORK, TO BE DISCUSSED LATER ON IN THE SEMESTER. IN ANY CASE, CHOMSKY'S IS A GENERAL POINT, WHICH ALL OTHER THINGS BEING EQUAL CERTAINLY HOLDS (THAT IS, DO WE EVER HAVE NON-INTERPRETABLE AGREEMENT FEATURES, AND IF SO WHAT DOES THAT MEAN?)

See Reid (1991) for discussion from a "functionalist" perspective. }

External manifestation of inflectional features appears to be the locus of much of the variety of languages, a topic that has gained prominence within the P&P framework.

THIS, OF COURSE, IS CRUCIAL, AND IN A SENSE A THIRD IMPERFECTION TO BE ADDED TO (34), PARTICULARLY IF, AS SEEMS PLAUSIBLE, THIS IS THE ONLY SOURCE OF CORE PARAMETRIC VARIATION. THE QUESTION THEN IS WHETHER ALL THE PROPERTIES IN (34), INCLUDING VARIATION--ALL OF THEM NON-OPTIMAL--COULD BE ACCOUNTED FOR IN SIMILAR OR IDENTICAL TERMS.

These observations presuppose that occurrences of features can be distinguished, a question raised earlier but put off: occurrences of agreement features are distinguished by category, some interpretable, some not. That falls well short of truly identifying occurrences. To do so would be necessary if feature chains exist.

THIS IS AN IMPORTANT ISSUE, AND WILL INTRODUCE AN CHANGE FROM CHAPTER 4. THE QUESTION IS AS SIMPLE AS IT IS DEEP. YOU HAVE SOME FEATURE F MANIFESTING ITSELF IN SOME CATEGORY K, AND A DIFFERENT FEATURE F' MANIFESTING ITSELF IN K'. ARE THEY OCCURRENCES OF THE SAME NUMERATION TOKEN, OR ARE THEY DIFFERENT TOKENS, OR A DIFFERENT POSSIBILITY EVEN, ARE THEY PROPERTIES THAT MANIFEST THEMSELVES IN CATEGORIES, EACH A DIFFERENT NUMERATION TOKEN? (INCIDENTALLY, YOU COULD ASK SIMILAR TOUGH QUESTIONS ABOUT CATEGORIES THEMSELVES, BUT THERE AT LEAST YOU HAVE THE PF COMPONENT--PERHAPS--TO MAP OCCURRENCES TO TOKENS; THERE ARE SOME COMPLICATIONS HERE AS WELL, BUT UNLESS SOMEBODY POINTS THEM OUT, I'LL SIDESTEP THEM FOR NOW). WHERE THE QUESTION MATTERS HERE IS IN TERMS OF WHETHER THERE ARE TRUE FEATURE CHAINS, A CENTRAL PART OF CHAPTER 4. OF COURSE, THE FACT THAT YOU POSSIBLY PRONOUNCE SOME FEATURES IN MORE THAN ONE PLACE ALREADY RAISES A WARNING FLAG, FOR AFTER ALL STANDARD CHAINS INVOLVE PRONUNCIATION IN A *SINGLE* SPOT (I'M THINKING, FOR INSTANCE, OF SO CALLED 'PARTIAL' WH-MOVEMENT).

THAT WORRY, THOUGH, IS NOT TOTALLY CONCLUSIVE, IF YOU BELIEVE NUNES'S LINE THAT THE REASON WHY YOU PRONOUNCE ONLY THE HEAD OF A CHAIN HAS TO DO WITH THE FACT THAT DIFFERENT COPIES (OCCURRENCES) OF THE SAME ITEM COULD NOT BE LINEARIZED, SINCE THEY WOULD COMMAND AND BE COMMANDED BY THE SAME ITEMS, THUS NOT YIELDING A LINEAR ORDERING IN KAYNE'S TERMS--SO YOU HAVE TO DELETE EVERY OCCURRENCE BUT ONE. (CHOMSKY WILL CHALLENGE THE LCA, SO HE WON'T BE MOVED BY THIS, BUT LET'S PUT THAT ASIDE NOW.) HOWEVER, NUNES ARGUES THAT WHICHEVER COPY HAPPENS TO MANAGE ITS WAY INTO THE MORPHOLOGICAL DOMAIN OF SOME HEAD X WILL NOT FACE ANY LINEARIZATION PROBLEMS, ASSUMING THAT WORD-LEVEL (MORPHOLOGICAL) RELATIONS ARE NOT DETERMINED BY THE LCA. IF YOU FOLLOW ME ALL THE WAY HERE, THEN YOU SHOULD BE ABLE TO ACCEPT THAT PRECISELY FEATURE CHAINS COULD HAVE SEVERAL PLACES OF PRONUNCIATION. BUT THERE ARE TWO CRUCIAL

ASPECTS TO NUNES'S REASONING, ONE OF THEM ARGUABLY A FORMAL TRICK. ONE IS THIS BUSINESS ABOUT MORPHOLOGICALLY ORDERED ITEMS NOT MATTERING, ACTUALLY CHOMSKY'S IDEA. THE REASON THIS IS PLAUSIBLE IS THAT, AFTER ALL, WORD-LEVEL RULES SEEM DIFFERENT; HOWEVER, IF WE GO BY LEVELS, IT IS TRICKY FROM NUNES'S PERSPECTIVE THAT LINEARIZATION CONCERNS OF THE PHASE-SYNTAX ALSO AFFECT HIGHER ORDERS, AS WE WILL SEE WHEN WE DISCUSS PARASITIC GAPS. PERHAPS THESE HIGHER ORDERS ARE ALSO SYNTACTIC, BUT THERE IS AN OBVIOUS WORRY HERE: WHY IS (ONLY) THE SYNTACTIC OBJECT (AT HIGHER ORDER THAN NUMERATION PHASES) SENSITIVE TO COMMAND? BUT THIS IS NOT WHAT CONCERNS ME NOW; MORE PROBLEMATIC IS THIS: WHAT DOES IT MEAN TO SAY THAT TWO COPIES ARE THE SAME? WE KNOW WHAT THIS MEANS FOR CHOMSKY: IN ESSENCE, CHOMSKY IS USING A QUANTUM OBJECT, WHICH APPEARS IN SEVERAL PLACES AT THE SAME TIME, BUT IS REALLY A SINGLE OBJECT. BUT THIS IS NOT WHAT NUNES IS SAYING AT ALL. FOR HIM EVERY COPY IS AS REAL AS IT GETS, AND IN FACT THIS IS SO MUCH SO THAT IN PARASITIC GAP CONSTRUCTIONS DIFFERENT COPIES OF THE SAME ITEM, WHICH HAPPEN NOT TO FALL UNDER A COMMAND PATH, MANAGE TO OBTAIN DIFFERENT REFERENTS AT LF. IN OTHER WORDS, FOR NUNES COPIES END UP BEING OCCURRENCES OR TOKENS DEPENDING ON WHETHER THEY FALL INTO A COMMAND PATH. OF COURSE, THIS IS FASCINATING, BUT WE SHOULD WORRY ABOUT WHETHER IT IS JUSTIFIED. THE KEY QUESTION IS THIS: TO SAY THAT SOMETHING IS THE SAME AS A COPY OF ITSELF JUST BECAUSE IT IS WITHIN THE SAME COMMAND PATH IS NECESSARY? THAT IS, COULDN'T IT BE OTHERWISE? WHAT IF THEY TOLD US THAT EXACTLY THOSE COPIES THAT ENTER INTO A COMMAND PATH ARE *NOT* OCCURRENCES OF THE SAME TOKEN? WOULD THAT BE INCOHERENT? I MUST CONFESS THAT I DO NOT REALLY SEE WHY THAT STATE OF AFFAIRS WOULD BE INCOHERENT, AND IF I'M RIGHT ABOUT THIS WORRY, THEN IT DOESN'T FOLLOW THAT THIS IS A FORMAL PROPERTY OF THE SYSTEM, AND THEN IT IS A SUBSTANTIVE PROPERTY THAT IS NOT JUSTIFIED WITHIN MINIMALISM, SINCE IT IS NOT SOMETHING WE CAN BLAME ON INTERFACE CONDITIONS--AT LEAST I DON'T SEE HOW WE CAN. IN ANY CASE, THESE ISSUES WILL BE AGAIN DISCUSSED WHEN WE RETURN TO PARASITIC GAPS, WHICH NUNES HAS A GREAT LINE FOR BUT WHICH WOULD SEEM TO FACE THIS CONCEPTUAL DIFFICULTY. FOR NOW ALL I HAVE TO SAY IS THAT, IF NUNES'S GENERAL REASONING FAILS, THEN IT IS SURPRISING THAT DIFFERENT FEATURES SHOULD BE PRONOUNCED IN DIFFERENT PLACES AT PF--SUGGESTING THE RELEVANT FEATURES SIMPLY DO NOT FORM A CHAIN.

{A possible case is control contingent on long-distance agreement.

THAT IS, WE'RE DEALING WITH INSTANCES LIKE *there arrived two men without PRO identifying themselves*, AND SIMILAR EXAMPLES INVOLVING PRO IN THE ADJUNCT, AND THEN A REFLEXIVE PRESUMABLY BOUND TO PRO.

See <MP>, chap. 4.5 (examples (40)-(43)), reviewing work of Anna Cardinaletti and Michal

Starke. It was assumed there that both binding and control are subject to these long-distance effects, but closer examination of Italian by Carl Cecchetto (pc) indicates that the effects are limited to control alone. That would explain the failure of binding in such examples as "there seem to each other [to be many men in the room],"

THIS IS AN OBVIOUS INSTANCE, ALTHOUGH IT INVOLVES MORE COMPLEX ISSUES (THE BENEFACTIVE CONFOUNDS THEM). THERE WAS AN INTERESTING WEB DISCUSSION ABOUT THIS A COUPLE OF YEARS AGO, MOSTLY BY PESETSKY AND LASNIK, WHICH ACTUALLY STARTED WITH A LITTLE EXAMPLE IN MY THESIS: there arrived two knights on each other's horses. AT THE TIME (MORE THAN A DECADE AGO, REALLY), I WAS TRYING TO ESTABLISH THAT BINDING THEORY AS UNDERSTOOD THEN WAS ESSENTIALLY A HAPHAZARD COLLECTION OF STATEMENTS. THIS EXAMPLE WAS TRYING TO SHOW THAT CONDITION A HOLDS AFTER EXPLETIVE REPLACEMENT, SINCE BY THE ASSUMPTIONS I WAS MAKING THE ANTECEDENT DID NOT COMMAND THE ANAPHOR PRIOR TO THAT.

WITH CURRENT ASSUMPTIONS, LASNIK ACTUALLY MANAGED TO DEMOLISH MY ARGUMENT IN THAT WEB DISCUSSION, BY OBSERVING THAT THIS IS BAD: *The DA proved [there to have been two men at the scene] during each other's trials. NOTE THAT, GIVEN THAT two men SHOULD RAISE NOT TO THE EXPLETIVE HERE, BUT HIGHER UP, TO AgrO OR v, TO CHECK ITS CASE, IT SHOULD BE ABLE TO BIND THE ANAPHOR FROM UP THERE. (YOU COULD TRY SOMETHING HERE, BUT GOOD LUCK: PERHAPS THE NEED TO MOVE BOTH THE EXPLETIVE AND each other CREATES A BINDING CONFLICT, WITHOUT COMMAND OBTAINING BETWEEN THE MOVED two men, THE MOVED there, AND THE MOVED ANAPHOR.) ANYWAY, THERE IS AN INTERESTING, RECENT LI SQUIB BY CARDINALETTI DISCUSSING THESE SORTS OF EXAMPLES, AND THE BOTTOM LINE IS THIS: IT APPEARS CLEAR THAT (OBLIGATORY) CONTROL, FOR SOME REASON, CAN BE SATISFIED IN A BROADER SET OF CASES THAN (STANDARD ANAPHORIC) BINDING CAN. I HAD INSTANCES IN R&R OF THE FOLLOWING SORT: There are two heavweights in tonight's bout because of each other's challenge, WHERE THERE OBVIOUSLY IS NO CONTROL ISSUE. HOWEVER, I HAVEN'T BEEN ABLE TO PERSUADE ANYONE--EVEN MYSELF--THAT THESE ARE GOOD (ACTUALLY, AS OPPOSED TO SOMETHING LIKE there's two heavweights in tonight's bout because of each other's challenge, WITHOUT AGREEMENT AND THUS, PLAUSIBLY ANYWAY (PARTICULARLY IF CARDINALETTI WAS GENERALLY RIGHT) WITHOUT EXPLETIVE REPLACEMENT BY two heavweights.

and fits with other evidence that binding requires an "overt" (possibly PRO, trace, or pro) antecedent, not just an implicit argument; see Rizzi (1986).

BUT WHAT DOES IT MEAN FOR A SYNTACTIC PHENOMENON TO HAVE TO BE OVERT? ACTUALLY, I THINK I KNOW ANOTHER ONE JUST LIKE THAT: TOPICALIZATION. THERE IS OVERT FOCUS AND (PLAUSIBLY) LF FOCUS, AND OVERT WH-MOVEMENT AND (PLAUSIBLY) LF WH-MOVEMENT. BUT I HAVEN'T SEEN A SINGLE CONVINCING CASE OF COVERT TOPICALIZATION. (I WAS CRAZY ENOUGH IN THE LAST CHAPTER OF MY THESIS--CHAPTER 5 OF THE PUBLISHED

VERSION--TO TRY AND ARGUE FOR THIS, CLAIMING THAT A DE-RE INTERPRETATION OF A LOWER, SAY, NAME, HAD TO INVOLVE SCOPING OUT TO A MATRIX POSITION, WHICH WAS TAKEN TO BE LF TOPICALIZATION; I DON'T THINK I CONVINCED ANYONE, AGAIN INCLUDING MYSELF.) SO IF TOPICALIZATION--THE CREATION OF AN ANTECEDENT--AND BINDING--THE RELATION TO AN ANTECEDENT--MAY NEED TO BE DONE OVERTLY, OR MORE PRECISELY NOT BY MOVEMENT, WE'RE MISSING SOMETHING BIG TIME ABOUT ANTECEDENCE. PERHAPS THE RELATION INVOLVES A COMPONENT WHICH RELIES ON INFORMATION-THEORETIC CONSIDERATIONS, WHICH MAKES IT, FOR REASONS OF SALIENCE (WHATEVER THAT MEANS) NECESSARILY OVERT.

There are additional complications. The phenomena seem less clear for passives than unaccusatives (possibly because of interference from an implicit subject), and become blurred or disappear in "long-distance" cases; see Hornstein (1996), Lasnik (1997).}

REMEMBER, THIS WAS ALL TRYING TO ESTABLISH FEATURE CHAINS PROPER. THE LOGIC WAS THIS: IF YOU MANAGE TO SAY THAT A FEATURE IS AT LF IN A POSITION DIFFERENT FROM WHERE IT IS AT SPELL-OUT, THEN THAT'S A CHAIN, BY ANY REASONABLE MEASURE. THAT'S WHY THE BINDING FACTS WOULD HAVE BEEN IMPORTANT. OF COURSE, WHETHER THE CONTROL FACTS HOLD MAY ALSO BE IMPORTANT, BUT THAT DEPENDS ON A MORASS OF ISSUES THAT I DON'T WANT TO GO INTO NOW. SO LET'S ASSUME THE GENERAL IDEA AND DENY THE EXISTENCE OF FEATURE CHAINS--IN EFFECT DENYING ATTRACT--AND SEE WHAT HAPPENS. EXPLICITLY:

In the absence of clear evidence to the contrary, I will assume that feature chains do not exist, hence that features cannot move or be attracted.

{Contrary to <MP>. The extension of these notions to features raises difficulties, not insuperable but better avoided, as seems possible.}

I SUPPOSE these notions REFERS TO MOVE AND ATTRACT. IT WAS ALWAYS UNCLEAR WHY, FOR INSTANCE, FEATURES COULD BE ATTRACTED IN MP, BUT THEY COULD NOT PROJECT THEIR OWN STRUCTURE, ASIDE FROM THE OTHER DIFFICULTIES MENTIONED ABOVE.

The dislocation property (II) is another apparent imperfection. In (35), for example, the phrase "an unpopular candidate" is in the natural position of interpretation as object of "elect" in (i)-(ii), but not in (iii)-(iv), though the interpretation is in relevant respects the same:

- (35)
- (i) they [elected an unpopular candidate]
 - (ii) there was [elected an unpopular candidate]
 - (iii) an unpopular candidate was elected
 - (iv) there was an unpopular candidate elected

AS I THINK I SAID BEFORE, THIS IS AN IMPORTANT PARADIGM THAT I WANT TO RETURN TO LATER ON IN THE SEMESTER (IN PARTICULAR THE CONTRAST BETWEEN (ii) AND (iv).)

In (iii)-(iv), the surface phonetic relations are dissociated from the semantic ones.

{We continue to disregard (iv), abstracting to the expected form (ii); see note <PROB>.}

THIS IS WHAT WE *WON'T* DISREGARD; THE REFERENCE IS TO FOOTNOTE 40, WHICH I THINK I ALREADY COMMENTED ON. MORE LATER ON.

Such phenomena are pervasive. They have to be accommodated by some device in any adequate theory of language, whether it is called "transformational" or something else.

Dislocation of $_$ yields a chain ($_$, $\langle t \rangle$) -- more accurately, a chain $\{X, Y\}$, where X and Y are occurrences of $_$. The raised element typically c-commands its trace in the original position, but where true, that follows from independent properties of C<HL>. Further operations might lead to violation of c-command and of locality relations between the two positions, as in multiple head-raising or independent XP-dislocation:

(36)

(i) $[[V\langle j \rangle - T]\langle i \rangle - C] [\dots \langle t \rangle \langle i \rangle \dots [VP \dots \langle t \rangle \langle j \rangle \dots]]$

(ii) [written $\langle t \rangle \langle j \rangle$ for children] $\langle i \rangle$, [those books] $\langle j \rangle$ couldn't possibly be $\langle t \rangle \langle i \rangle$]

NOTE THAT WHETHER OR NOT COMMAND OBTAINS IN (i) DEPENDS ON TRICKY ISSUES CONCERNING ADJUNCTION.

In (36), the indices are redundant, the chains determined by constitution of the trace. That need not be so, however, as in (37):

(37) whom $\langle i \rangle$ did everyone talk to whom $\langle j \rangle$ about whom $\langle k \rangle$

THIS IS A VERY CONFUSING EXAMPLE. OF COURSE, (37) IS NOT A PF. THE POINT HERE IS THAT YOU CAN FORM CHAINS FROM EITHER THE *to* PHRASE OR THE *about* PHRASE POSITIONS, AND MERE COMPUTATIONAL CONDITIONS DO NOT DICTATE THE CHOICE.

Chains can be formed with $\langle i \rangle = \langle j \rangle$ or $\langle i \rangle = \langle k \rangle$, both consistent with locality conditions; and the full range of interpretations seems to be available, either way.

In the approach we are pursuing here, the chains at LF are determined by identity throughout, with the ambiguity of (37) resolved by the derivation, given the initial numeration.

THAT IS, THE AMBIGUITY DOESN'T REALLY ARISE IN THE DERIVATIONAL

SYSTEM. YOU HAVE TWO who'S IN THE NUMERATION AND ONE MOVES CREATING A CHAIN, AND UNDER IDENTITY OF OCCURRENCES YOU IDENTIFY THAT CHAIN, THUS RECOVERING THE NUMERATION TOKEN. BUT REMEMBER, THIS IS A WAY OF EXECUTING THE PROGRAM, NOT A NECESSARY ONE. IF YOU'RE A REPRESENTATIONALIST, YOU WANT TO FIND WAYS OF FORCING THE RELEVANT PF AND LF INTERPRETATIONS OF (37), WITHOUT RESORTING TO NUMERATIONS AND SO ON. AND IF YOU DON'T WANT THE NUMERATION, THEN:

Other means would be required if we were to eliminate this device along lines discussed earlier. In a strict derivational approach semantic interpretation is cyclic and the problem of interpreting (36), (37) arises in a somewhat different form. In a representational approaches chains are determined by an algorithm A operating in a "search space" of possible LFs; the burden of accounting for locality and other conditions on chains would then rest on A. Here we return to issues of general architecture raised in sec. 3.

SEE FOR INSTANCE BRODY'S MONOGRAPH FOR THIS SORT OF APPROACH, OR MUCH OF KOSTER'S WORK (INCLUDING EARLIER ONE) AND PART OF RIZZI'S.

However these matters are resolved, we have two "imperfections" to consider: uninterpretable features and the dislocation property.

AGAIN, UNDERSTOOD IN TERMS OF ALGORITHM A IF YOU CHOOSE NOT TO DO THINGS DERIVATIONALLY. THE PROBLEM DOESN'T GO AWAY, THAT IS.

These properties (in fact, morphology altogether) are never built into special-purpose symbolic systems. We might suspect, then, that they have to do with externally-imposed legibility conditions.

I UNDERSTAND THE TWO CLAIMS, BUT I DON'T REALLY SEE HOW THE then RELATES THEM. THAT IS, I DON'T SEE HOW ANYTHING WE SAY ABOUT SYMBOLIC SYSTEMS REALLY BEARS ON THESE ISSUES. IN ANY CASE, LET'S GO WITH THE GENERAL IDEA, WHICH SEEMS IF NOT PLAUSIBLE (WHO KNOWS) AT LEAST INTERESTING.

With regard to dislocation, that has been suggested from the earliest days of modern generative grammar, with speculations about facilitation of processing (on the sound side) and the dissociation of "deep" and "surface" interpretive principles (on the meaning side). The boundaries are not clear,

{They have also shifted as inquiry proceeded: thus from the 1950s through the mid-'70s, such matters as quantifier scope were commonly taken to be prototypical examples of "surface" interpretation, while more recent work generally takes them to be prototypical properties at LF.}

nor the mechanisms to express them. One approach to the array of problems was to distinguish the role of deep and surface (D- and S-) structure in semantic interpretation: the former enters into determining quasi-logical properties such as entailment and theta structure; the latter such

properties such as topic-comment, presupposition, focus, specificity, new/old information, agentive force, and others that are often considered more discourse-oriented, and appear to involve the "edge" of constructions.

OF COURSE, THIS PARTICULAR CUT WAS BEFORE LF WAS SYSTEMATICALLY STUDIED. I DON'T THINK IT IS REALLY EASY TO ARGUE THAT SURFACE SYNTAX COVERS ALL MATTERS PERTAINING TO SPECIFICITY, FOCUS, AND THE LIKE, UNLESS ONE IS READY TO CLAIM THAT ALL SUCH MATTERS INVOLVE, SAY, THE DISPLACEMENT ONE SEES IN HUNGARIAN OR IN GERMAN *AT SURFACE SYNTAX*. OF COURSE, THAT'S PRECISELY WHAT KAYNE HAS BEEN ARGUING RECENTLY, A MATTER THAT I PLAN TO RETURN TO AFTER WE INTRODUCE MANY MORE ASSUMPTIONS IN THE COURSE. BUT IF YOU SET THAT ASIDE, THERE JUST SEEMS TO BE THE CASE THAT IN LANGUAGES LIKE ENGLISH YOU DO NOT OVERTLY FOCALIZE AND YOUR SPECIFIC EXPRESSIONS CAN BE LEFT IN SITU, AND NEED NOT BE OVERTLY SCRAMBLED. AT ANY RATE, I DO THINK THE WORLD WOULD BE A NICER PLACE IF IT WERE TRUE THAT SURFACE SYNTAX DOES DO ALL OF THIS, HAVING TO DO WITH INFORMATION STRUCTURE, AND NONE OF THESE ISSUES HAPPEN AT LF. BUT THAT POSES A FURTHER QUESTION: WHY? ARE THERE, THEN, ANY LF OPERATIONS? ONE VIEW, NOTABLY EXPLORED BY BEGHELLI, FOX, REINHART, AND OTHERS, WOULD ARGUE THERE ARE (WE SHOULD SEE IF THEY HAVE CONVINCING CASES); AS I SAID KAYNE ARGUES THERE AREN'T, RADICALLY SO EVEN FOR TRADITIONAL QUANTIFIER RAISING INSTANCES. THAT COULD BE, BUT POSES TWO NEW QUESTIONS: (I) ARE THERE ANY IN-PRINCIPLE DIFFERENCES BETWEEN INFORMATION-RELATED RELATIONS (SPECIFICITY AND SO ON) AND INTENTIONALITY-RELATED RELATIONS BEARING ON TRUTH CONDITIONS? (II) MORE GENERALLY, IS THE LF COMPONENT ACTIVE AT ALL, OR DOES IT SIMPLY PIGGY-BACK ON WHATEVER HAPPENS FIRST, ALONG THE LINES OF KITAHARA/HORNSTEIN? NEEDLESS TO SAY, THESE ARE ALL DEEP ISSUES.

AND HERE'S THE FINAL ONE: SHOULD WE CONFLATE THOSE BASIC RELATIONS (SUCH AS THETA-STRUCTURE AND LEXICAL ENTAILMENT) THAT USED TO BE THOUGHT OF AS THE DOMAIN OF D-STRUCTURE WITH RELATIONS THAT TAKE PLACE AT LF? THAT INFERENCE IS INVITED BY GETTING RID OF D-STRUCTURE, BUT IN THE LIMIT IT WILL JUSTIFY GETTING RID OF THOSE NOTIONS ALTOGETHER (THETA STRUCTURE, LEXICAL ENTAILMENTS)--HOWEVER, IS THAT A GOOD IDEA? A RADICAL WAY OF PUTTING THIS, WITHIN THE PRESENT CONTEXT ASKING ABOUT DISPLACEMENT, IS THIS: IF IT'S ALL BUNDLED UP AT LF, WHY DON'T WE HAVE LEXICAL CHOICES FOR SUCH THINGS AS DIFFERENT SCOPAL ORDERS OR DIFFERENT FOCAL PROPERTIES? IN OTHER WORDS, WHY ISN'T THERE A LANGUAGE THAT ALLOWS YOU TO SAY someone glibs everyone TO MEAN THAT EVERYONE IS SUCH THAT HE OR SHE LOVES SOMEONE, AND someone bligs everyone TO MEAN THAT SOMEONE IS SUCH THAT EVERYONE LOVES HIM OR HER? OR WHY CAN'T WE FIND A LANGUAGE THAT ALLOWS YOU TO SAY John babum Mary TO MEAN THAT JOHN LOVES *MARY* AND John badabum Mary TO MEAN THAT *JOHN* LOVES MARY. THOSE LANGUAGES WOULD BUILD WORDS

NOT AROUND THETA STRUCTURE AND ENTAILMENT RELATIONS, LIKE OURS DO, BUT IN TERMS OF INTENTIONAL AND INFORMATION-DRIVEN RELATIONS (OR AT ANY RATE, BOTH). IN AS MUCH AS NONE OF THAT EXISTS--ALTHOUGH IT IS PERFECTLY IMAGINABLE AND SENSIBLE AS A FORMAL SYSTEM--THERE PLAINLY IS A MISMATCH BETWEEN THE DIFFERENT KINDS OF RELATIONS (SAY, BETWEEN CONCEPTUAL STRUCTURE AND INTENTIONAL STRUCTURE). THAT'S THE MYSTERY.

Theories of LF and other approaches sought to capture the distinctions in other ways. The "deep" (LF) properties are of the general kind found in language-like systems; the "surface" properties appear to be specific to human language.

MORE PRECISELY, THE "DEEP" PROPERTIES WERE INTENTIONAL (ULTIMATELY TRUTH-FUNCTIONAL, WHICH NEED NOT BE, INCIDENTALLY, TRUTH CONDITIONAL OR EVEN TRUTH DESAMBIGUATING--SEE MAY (1985)), WHEREAS THE "SURFACE" STRUCTURE WERE NOT (E.G. THE RELATIVE HEIGHT OF SPECIFIC ELEMENTS DOESN'T CHANGE TRUTH CONDITIONS). WHETHER BETWEEN NON-INTENTIONAL PROPERTIES YOU MAY ALSO SEPARATE A LAYER OF CONCEPTUAL PROPERTIES DEPENDS ON SUCH ISSUES AS WHETHER THERE IS SPECIFIC THETA INFORMATION (HIERARCHIES, THETA CONDITIONS OF THE DOWTIAN SORT, PERHAPS BAKER'S UTAH, ETC.).

If the distinction is real, we would expect to find that language design marks it in some systematic way -- perhaps by the dislocation property, at least in part.

THAT'S A BIT OF WISHFUL THINKING. THE REAL DISTINCTION CAN BE MARKED IN A ZILLION WAYS, AND MORE IMPORTANTLY, IT NEED NOT BE MARKED AT ALL... ON THE OTHER HAND, IF MOVEMENT WERE AN IMPERFECTION (I.E. YOU BUY THE BULLET), THEN YOU COULD RUN THE ARGUMENT THAT THE HUMAN MIND EXAPTED IT FOR "MARKING" (I.E., IN CHOMSKY'S SENSE, PRESUMABLY FOR PURPOSES OF SALIENCE IN THE MAPPING WE'RE NOW CONSIDERING), JUST AS IT PRESUMABLY EXAPTED A WEIRD PROPERTY OF LARYNXES FOR CODING NUMERATION TOKENS. SOMETIMES ACCIDENTS ARE NICE, AND OUR UNIVERSE IS THERE TO PROVE IT! JUST LOOK AT HOW USEFUL YOUR THUMB HAS TURNED OUT TO BE... THE POINT IS, THE STRONG MINIMALIST THESIS, AS FAR AS I CAN SEE, WOULDN'T SUFFER IF IT IS TRUE ONLY TO THE EXTENT THAT IT IS, AND SOME ACCIDENTAL PROPERTIES OF PF/MORPHOLOGY LEAD TO CURIOUS RESULTS, LIKE WORD TOKENS OR CHAINS. OF COURSE, WITHOUT THE STRONG MACHINERY BEHIND IT, ALL OF THAT WOULD LEAD, AT BEST, TO A BIRD-SONG! BUT THE FACT THAT YOU'RE ARGUING FOR A WONDERFULLY PRETTY MACHINE DOESN'T MEAN THAT IT CANNOT HAVE PATCHWORK PARTS, EVEN PATCHWORK PARTS THAT MAKE IT USABLE. AFTER ALL, NOTHING IN THE PRETTY DESIGN HAD ANYTHING TO DO WITH USABILITY, LEARNABILITY, AND ALL THE REST! (AND INSPITE OF THE METAPHORS, THERE *WASN'T* AN ACTUAL ENGINEER OPERATING HERE). ANYWAY, NOT TO BLAME CHOMSKY OF THIS GOULDIAN RETREAT (?), CALL THIS URIAGEREKA'S "MIXED" STRONG THESIS. (YOU'LL SEE

THAT, IN SOME RESPECTS, I'LL PRESENT A THESIS THAT IT'S EVEN STRONGER THAN CHOMSKY'S, BUT I HAVE NO SCRUPLES IN MAKING IT DIRTY AS WELL, PRECISELY BECAUSE WE'RE DEALING WITH BIOLOGICAL, DIRTY BEINGS.)

To the extent that such ideas can be given substance, it would follow that the dislocation property is required; it falls within the design specifications given to the super-engineer seeking an optimal solution to conditions imposed by the external systems.

OF COURSE, WE CAN GO ALONG AND BUY THIS JUST TO SEE WHERE IT LEADS, BUT I MUST SAY I DON'T FOLLOW THE PUNCH-LINE. WE MAY AS WELL CALL THIS THE "PRINCIPLE OF HOPE"... (GOOD ONE TO HAVE, THOUGH.)

This line of argument might provide motivation for the dislocation property, but it would remain to find the mechanisms employed to implement it. The distinction is familiar. We may say that the function of the eye is to see, but it remains to determine the implementation; a particular protein in the lens that refracts light, etc.

ACTUALLY, THAT'S THE KIND OF EXAMPLE THAT ADDS KINDLE TO MY FIRE ABOVE. WHAT DOES IT MEAN TO SAY THAT THE EYE'S *FUNCTION* IS TO SEE? THAT SEEMS LIKE AN A POSTERIORI COMMENT. MORE LIKELY, EYES ARE SO COMMON (THEY INDEPENDENTLY EVOLVED DOZENS OF TIMES, THOUGH MORE ON THIS SHORTLY) BECAUSE OF A COUPLE OF FACTORS. FIRST, BECAUSE LIGHT SENSITIVE CELLS (WITH THE RELEVANT PROTEIN, ETC.) DON'T SEEM PARTICULARLY HARD TO COME BY, PERHAPS NOT SURPRISINGLY IN A WORLD LIKE OURS THAT GETS LOADS OF LIGHT, APPROPRIATELY POLARIZED (OTHERWISE IT MIGHT BE TOO STRONG TO USE FOR FORMING IMAGES). SECOND, BECAUSE THE SORT OF INVAGINATION YOU NEED TO CAPTURE IMAGE-FORMING EYES IS SOMETHING THAT MANY COMPLEX ORGANISMS NEED ANYWAY. TO BEGIN WITH, WE START AS AN INVAGINATED TUBE IN THE PRIMORDIAL AXIS, AND FURTHERMORE MOST FEMALES (THOUGH NOT ALL, CURIOUSLY) INVOLVE SOME SORT OF INVAGINATION FOR REPRODUCTION. AS IT TURNS OUT, THE GENES RESPONSIBLE FOR VULVA-FORMATION IN MANY SPECIES ARE ALSO RESPONSIBLE FOR CORE EYE-FORMATION. THAT, I'D SAY, IS A PRETTY LUCKY ACCIDENT; IF WE HADN'T FOUND MULTICELLULAR ORGANISMS OR EVEN SEX, WE WOULDN'T HAVE FOUND EYES. PUT DIFFERENTLY, IN THE WORLD OF BACTERIA (WHICH IS MOST OF THIS PLANET, NOW AND ALWAYS) THERE ARE NO EYES, ALTHOUGH THERE PROBABLY WOULD BE A PRETTY NICE FUNCTION FOR EYES. (IN THE WORLD OF EUKARIOTIC BEINGS, INCLUDING UNICELLULAR EYES, APPARENTLY THERE ARE ALREADY SOME LIGHT SENSITIVE DEVICES IN THE CYTOSKELETON, BUT IT'S HARD TO CALL THAT AN EYE--CERTAINLY, IT ISN'T IMAGE-FORMING, LET ALONE IMAGE-INTERPRETING!) BUT THERE ARE PROBABLY EVEN MORE ACCIDENTS BEHIND THE FORMATION OF IMAGE-FORMING EYES. AS IT TURNS OUT, THE GENE RESPONSIBLE FOR DIRECTING THE DESIGN OF SUCH EYES (APARENTLY DICTATING SUCH THINGS AS THE SPOT OF AN ALREADY POSSIBLE INVAGINATION, BUT WITH LIGHT SENSITIVE CELLS AND A CERTAIN CHUNK OF TRANSPARENT CELLS TO FOCUS THE IMAGE--WITHOUT

WHICH YOU'D ONLY HAVE A WIDE ANGLE PINHOLE-STYLE CAMEREA OBSCURA) IS PRESENT IN AN ANCESTOR COMMON TO MICE, FLIES, SQUID, AND GOD KNOWS WHAT ELSE--THE SO-CALLED EYELESS GENE. THAT'S PRETTY FAR BACK, AND PRESUMABLY ANOTHER ACCIDENT (IN OTHER WORDS, PART OF THE SURPRISING CONVERGENCE BETWEEN, SAY, A SQUID EYE AND AN INDEPENDENTLY EVOLVED FLY EYE IS NOT SO SURPRISING). THE POINT IS, THERE IS NO FUNCTION TO ANY OF THIS UNTIL THE STRUCTURES ARE IN PLACE, PARTLY BECAUSE OF ACCIDENTS AND PARTLY BECAUSE OF DEEP STRUCTURAL PROPERTIES OF ORGANISMS, SUCH AS THEIR HAVING TRANSLUCID CHUNKS OR LIGHT-SENSITIVE CHUNKS. THAT'S HOW FORMS EMERGE; WHETHER THEY *STAY* OF COURSE IS A DIFFERENT MATTER: THEY DON'T IF THEY'RE HARMFUL (MEANING THE ORGANISM DIES BEFORE REPRODUCTION), THEY DO IF THEY'RE NOT, AND THEY THRIVE IF THEY'RE ADAPTIVE, LEADING TO MUCH SEX AND MANY CHANGES AND SO ON. SO ANYWAY, IF THE EYE IS ANY INDICATION, MY MIXED STRONG THESIS CLEARLY APPLIES TO IT, IF YOU'RE GENEROUS WITH INTERPRETATION. SO THEN THE QUESTION IS WHETHER YOU WANT TO SAY THAT THE LINGUISTIC SYSTEM IS REALLY VERY DIFFERENT. OF COURSE, WHO KNOWS--HERE WE HAVE ONLY ONE, AND WE CAN'T REALLY TORTURE OURSELVES TO FIND OUT ALL THE THINGS WE'VE FOUND ABOUT EYES... BUT ANYWAY, PROFOUNDLY, THAT'S PARTLY WHAT'S AT ISSUE--AND WHY I'M NOT TOO BOTHERED WITH HAVING SOME GLITCHES IN THE SYSTEM THAT WE CAN USE, IN FACT EVEN MORE PRODUCTIVELY THAN THE GRAND STRUCTURAL DESIGN. INDEED, I'LL GO AS FAR AS TO CLAIMING THAT WITHOUT GLITCHES YOU WOULDN'T HAVE ANYTHING, SINCE IT IS GLITCHES THAT DRIVE SYSTEMS INTO EXISTENCE TO BEGIN WITH. PERFECT FORM IS DEAD; WORSE, IT DOESN'T EXIST (DEATH ALREADY APPLIES TO AN IMPERFECTION: LIFE). I KNOW I'M GETTING PHILOSOPHICAL, SO I'LL STOP SAYING THIS: THE ISSUE IS NOT PERFECTION, BUT QUASI PERFECTION. THERE'S THE RUB: ALMOST PERFECT, WHICH MEANS YOU NEED TO UNDERSTAND THE PERFECTION AND THEN SEE WHAT SETS IT IN MOTION.

Similarly, certain semantic properties may involve dislocated structures, but we want to discover the mechanisms that force dislocation. Minimalist intuitions lead us to look at the other major imperfection, the uninterpretable inflectional features. Perhaps these devices are used to yield the dislocation property.

I THINK THIS IS BACKWARDS. I'D TRY TO RUN THE ARGUMENT THE OTHER WAY AROUND: YOU HAVE THOSE STUPID DEVICES FOR NO GOOD REASON (FOSSILIZED MORPHOLOGY, SAY), BUT HEY: THEY SET THE SYSTEM IN MOTION--LITERALLY, A CATASTROPHE--SO THAT IT CAN GET RID OF THEM. ONCE THE SYSTEM IS IN MOTION THAT WAY, YOU CAN USE THOSE DYNAMICS FOR OTHER PURPOSES--YOU COOPT THEM. I TRY TO RUN THAT LINE IN CHAPTER 6 OF R&R, OF COURSE IN THE VOICE OF THE OTHER... AT ANY RATE, FOR MOST OF THE ANALYSIS BELOW, WHETHER YOU THINK IT IS THE CHICKEN (MOVEMENT) OR THE EGG (FEATURES), IT DOESN'T MATTER.

If so, then the two imperfections might reduce to one, the dislocation property. But the latter might itself be required by design specifications. That would be an optimal conclusion, falling under option (iii) of (29).

SO TO SUM UP, I AGREE WITH THE CONCLUSION THAT THE IMPERFECTION IS ONLY ONE, BUT DISAGREE WITH THE CONCLUSION, WHICH ONLY THE PRINCIPLE OF HOPE YIELDS, THAT THIS HAS TO DO WITH DESIGN SPECIFICATIONS. AND I CLAIM THAT THIS MAKES NO DIFFERENCE SO FAR, PARTICULARLY BECAUSE WE DON'T KNOW HOW THE PRINCIPLE OF HOPE WORKS.

To establish any such conclusion is no simple matter.

NOTE, THE MIXED STRONG THESIS HAS TO ESTABLISH THIS CONCLUSION TOO.

We are entering terrain that is mostly unexplored. One approach is suggested by the observation that for convergence, uninterpretable features must be deleted in the course of computation of LF. Consider the dislocated example (35iii), repeated here:

THE TONE OF THESE SENTENCES IS WORTH EMPHASIZING. THERE IS NOTHING OBVIOUS OR OBVIOUSLY RIGHT ABOUT THE OBSERVATION (NOTE, A MERE OBSERVATION) THAT YOU MUST GET RID OF CERTAIN FEATURES. THIS IS ALL TENTATIVE, MOSTLY UNEXPLORED STUFF...

(38) an unpopular candidate T-was elected <t>

There are three kinds of uninterpretable features in this structure:

THESE ARE THE *RELEVANT* FEATURES--THERE'S OTHERWISE MANY MORE...

(i) the agreement features of T (taking them as a unit, the set of $_$ -features), (ii) the EPP feature of T that requires "second Merge," and (iii) the structural Case feature of "an unpopular candidate." The $_$ -set (i) identifies T as a target of dislocation;

CAREFUL WITH THIS, WHICH WILL SOON MOVE FROM AN OBSERVATION TO A FUNDAMENTAL PROPERTY OF THE NOTION 'AGREE'.

the EPP-feature (ii) requires that something be merged in this position; the Case feature (iii) identifies "an unpopular candidate" as a candidate for such merger (hence dislocation).

AND HERE TOO BE CAREFUL; THE IDENTIFICATION IN QUESTION--LESS OF AN OBVIOUS, OR MERE OBSERVATION IN THIS INSTANCE, PARTICULARLY SINCE CASE DOESN'T ALWAYS IDENTIFY A WHOLE RELEVANT PHRASE ACROSS LANGUAGES--WILL TURN INTO A CENTRAL PROPERTY OF 'AGREE'.

Successful implementation of the operation erases all of the uninterpretable features, forming MLIs with a reduced set of features.

THAT IS, BY ASSUMPTION.

The approach is optimal, if indeed uninterpretable features are the mechanism for dislocation.

THE LOGIC IS: YOU NEED ALL OF THOSE THREE COMPONENTS TO MAKE MOVEMENT WORK (WHAT TO DO, WHERE TO GO, WHERE TO COME FROM, WHICH RECALLS THAT OLD PHILOSOPHICAL TRINITY OF WHO WE ARE, WHERE DO WE COME FROM, AND WHERE ARE WE GOING!) AND LO AND BEHOLD, YOU HAVE THEM ALL. THE TRUTH IS, THOUGH, THAT THE INVERTED LOGIC ALSO WORKS. THAT IS, SUPPOSE YOU HAD UNINTERPRETABLE FEATURES FOR SOME REASON, SAY A RESIDUE OF INCORPORATED MORPHOLOGY (IN ESSENCE, WHAT THEY USED TO CALL 'GRAMMATICALIZATION'). THE SYSTEM IDENTIFIES THIS 'VIRUS' AND HAS TO GET RID OF IT, FAST; FOR THAT, IT NEEDS A COMPATIBLE ELEMENT TO GO OVER WHERE THE VIRUS IS AND KILL IT--A SORT OF ANTIBODY. SAY YOU IDENTIFY THE ANTIBODY IN TERMS OF CASE AND THE VIRUS IS THE NASTY UNINTERPRETABLE FEATURE. AS FOR EPP, YOU ADD IT TO THE PROCESS JUST AS YOU DO IN CHOMSKY'S REASONING: BECAUSE IT'S THERE. INDEED, FROM THE PERSPECTIVE OF THE MIXED THESIS YOU'D SAY: ONCE THE ANTIBODY GOES TO THE VIRUS, THE IDENTIFIED (CASE MARKED) PHRASE IS IN AN EDGE SITE (THE EPP DOMAIN), WHERE IT HAD TO BE ANYWAY--FOR SOME REASON. WE WILL RETURN LATER ON TO WHAT THAT REASON COULD BE, BUT IT WORKS THE SAME ON BOTH TREATMENTS, THE ONE HOLDING MOVE AS MOTIVATING FEATURES AND THE ONE HOLDING FEATURES AS MOTIVATING MOVEMENT.

{The approach renders superfluous the intuitive motivation for pied-piping proposed in <MP>, in terms of PF crash.

SPECIFICALLY IN TERMS OF PF CRASH, WHICH DOESN'T MEAN THE GENERAL PIED-PIPING LINE COULDN'T WORK FOR OTHER REASONS, INCLUDING AT LF, AS IN FOX'S DISSERTATION. AND ALSO:

See Uriagereka (forthcoming a) for argument that some such device nevertheless operates in languages with rich morphology, with consequences for null subject, barriers, and Spell-Out.

WE'LL SEE THIS PAPER LATER ON IN THE SEMESTER--THE BASQUE ONE. IN ANY CASE, THE ISSUE IN LANGUAGES LIKE BASQUE IS WHETHER RICH (OR WHAT I CALLED 'HEAVY') MORPHOLOGY INVOLVES NOT JUST PIED-PIPING FOR MORPHOLOGICAL REASONS, BUT FURTHERMORE THE EMERGENCE OF A KIND OF BARRIER IN THE SYSTEM. LET'S SET THIS ASIDE FOR NOW.

For an approach somewhat similar to what is outlined below, from a partially different perspective, see Frampton and Guttman (1998).}

TECHNICALLY, THE FRAMPTON AND GUTTMAN PERSPECTIVE IS RATHER

DIFFERENT, BUT THE GENERAL LINE OF REASONING, INCLUDING VARIOUS SPECIFICS ABOUT COMPUTATIONAL COMPLEXITY, IS EXACTLY THE SAME. BUT AGAIN, LET'S PUT THIS TO THE SIDE.

Assuming so, let us look more closely. Suppose that the derivation has constructed the syntactic object (39), having merged T with the copula-headed phrase:

(39) T be elected an unpopular candidate

The new element T has uninterpretable features of two types: its \bar{u} -set and its selectional feature EPP. Like other selectional features, EPP seeks an XP to merge with the category it heads.

THIS IS RATHER MISLEADING TERMINOLOGY AND PERHAPS EVEN IMPLEMENTATION. OF COURSE, WE CAN CALL 'SELECTION' WHATEVER WE WANT, BUT USUALLY THIS IS A TERM USED FOR THETA-THEORY, AND THUS WE SAY, FOR INSTANCE, THAT A TRANSITIVE VERB SELECTS FOR A DIRECT OBJECT, OR A DITRANSITIVE FOR WHATEVER YOU THINK IT DOES. BUT IT IS REALLY NOT VERY PROFOUND TO SAY THAT A DETERMINER SELECTS FOR AN NP OR A TENSE FOR A VP; IN A SENSE, OF COURSE THEY DO. BUT THEY ALWAYS DO! DETERMINERS ARE NOT CLASSIFIED IN TERMS OF WHETHER THEY START THEIR DERIVATIONAL LIFE SELECTING ONE OR TWO NOUN PHRASES, FOR INSTANCE. THERE'S NO SELECTION ISSUE THERE--IT'S JUST THE GARDEN VARIETY COMBINATORIAL PROPERTY OF THESE ELEMENTS. I AT LEAST THINK OF SELECTION AS SOMETHING MORE SUBSTANTIVE, IN PART BECAUSE I TAKE THETA-THEORY RATHER SERIOUSLY, AND I DON'T SEE HOW DETERMINERS OR TENSES HAVE SERIOUS THEMATIC PROPERTIES. ANYWAY, WORSE STILL PERHAPS, WHAT T IS SUPPOSED TO SELECT IN ITS SPECIFIER IS SOMETHING THAT CAN GO THERE EITHER BY MERGE OR BY MOVE. AGAIN, CALL IT WHAT YOU WILL, BUT I THINK IT'S GOOD TO KEEP THE TERM SELECT FOR SOMETHING YOU DO UNDER MERGE, AGAIN AS IN THETA THEORY. OR IF YOU CALL SELECT THE MORE GENERAL THING, THEN WHATEVER HAPPENS UNDER MERGE WITH THETA ROLES AND SO ON IS NOT SELECT, BUT SOMETHING NARROWER. THE POINT IS: CALLING THE EPP PROBLEM A SELECTION FEATURE WON'T MAKE IT GO AWAY. WE REALLY HAVE NO CLUE AS TO WHAT'S GOING ON IN EPP SITES. THIS PAPER IS TRYING TO SAY SOMETHING MEANINGFUL ABOUT IT--AND I THINK IT DOES--BUT I WOULDN'T CALL IT A SELECTION ISSUE, SINCE IT CLOUDS THE REAL DIFFICULTY.

The \bar{u} -set we can think of as a <probe> that seeks a <goal>, namely, "matching" features that establish agreement. The relation of the probe of T to its goal is the T-<associate> relation.

OKAY, SO THIS IS NOW ALL NEW TERMINOLOGY, THOUGH INTUITIVE. THE IDEA IS THAT THE PROBE IS WHAT WE USED TO CALL IN THE OLD DAYS (LITERALLY IN THE SAME CONTEXTS IN THE R&U FRAMEWORK) THE 'GOVERNOR' AND OF COURSE THE 'GOAL' IS WHAT WE USED TO CALL THE 'GOVERNEE', AND EXACTLY WHERE WE USED TO SAY AGREEMENT CAN HOLD BY LICENSING OF THE

FEATURES OF THE GOVERNEE, NOW WE SAY THAT THERE IS A MATCHING OF FEATURES THAT APPEAR ON BOTH PROBE AND GOAL. I HOPE YOU SEE THIS IS TECHNICAL IMPLEMENTATION, STILL CODING A LONG-DISTANCE RELATION IN PRETTY MUCH THE SAME WAY IT USED TO BE CODED. INDEED, I SHOULD SAY THAT BY THE END OF THE PAPER THE RESEMBLANCE OF THE SYSTEM TO THE R&U PAPER IS EVEN MORE DRASTIC, SINCE IN THAT PAPER WE HAD LONG-DISTANCE GOVERNMENT APPLY OVERTLY PRECISELY TO DISTINGUISH IT FROM WHAT AT THE TIME WERE THOUGHT TO BE CATEGORIAL DISPLACEMENTS AT LF (IN THOSE DAYS THE ENTIRE ASSOCIATE, NOT JUST ITS FEATURES, WAS TAKEN TO MOVE TO THE TARGET), WHICH WE TRIED TO SHOW WERE NOT OPERATIVE IN THE EXAMPLES WE WERE ANALYZING. SO R&U WAS A THEORY OF AGREEMENT UNDER GOVERNMENT IN THE OVERT SYNTAX, WHERE GOVERNMENT WAS DEFINED IN TERMS OF A DYNAMIC NOTION THAT WAS PRACTICALLY IDENTICAL TO THE EFFECTS OF THE MINIMAL LINK CONDITION (INDEED, IN SOME LATER WORK WE CALLED THE NOTION 'LOCALITY' INSTEAD OF 'GOVERNMENT', ALTHOUGH THAT'S JUST TERMINOLOGY). AT THIS POINT IN THIS PAPER, IT SEEMS AS IF THE NOTION OF AGREEMENT BEING DISCUSSED IS NOT SOMETHING THAT TAKES PLACE IN OVERT SYNTAX, THUS IS DIFFERENT FROM WHAT WE HAD. BUT BY THE END OF THE PAPER, ONLY OVERT SYNTAX REMAINS! SO AGAIN, AS FAR AS I CAN SEE THE NOTIONS ARE FOR ALL INTENTS AND PURPOSES FORMALLY THE SAME.

For the $_set$ of T in (39), there is only one choice of matching features: the $_set$ of "candidate." Locating this goal, the probe erases under matching. Taking structural Case to be a reflex of an uninterpretable $_set$,

THE NOTION OF ERASURE UNDER MATCHING IS WHAT SOMETIMES GETS REFERRED TO AS 'SUICIDAL GREED'. THAT IS, THE PROBE F SEEKS AN IDENTICAL FEATURE F' IN THE GOAL, AND WHEN F FINDS F', F GOES, LITERALLY ERASES--KAPUT BIG TIME, NOT JUST SOME CODING SAYING 'DON'T INTERPRET ME AT THE INTERFACE,' AS IN ELLIPSIS SITES, SAY.

{If a reflex of an interpretable $_set$, it would be erased in situ by the $_set$ of <candidate> itself.}

I MUST CONFESS I DON'T GET THIS FOOTNOTE. IS CHOMSKY TALKING ABOUT INHERENT CASE, WHICH COULD REASONABLY BE SAID TO BE THE REFLEX OF SOMETHING INTERPRETABLE? OR IS HE TALKING ABOUT SOME OTHER FORM OF CASE? PART OF THE REASON I DON'T JUMP INTO CONCLUSIONS HERE IS BECAUSE OF THE PHENOMENON OF CASE SPREAD IN LANGUAGES LIKE RUSSIAN, WHERE IT IS NOT ENTIRELY OBVIOUS TO ME A) WHY CASE APPEARS IN SEVERAL PLACES WITHIN THE NOMINAL EXPRESSION, AND B) HOW DO YOU GET RID OF ALL OF THOSE?--OR CAN YOU SAY THAT THOSE ARE INTERPRETABLE?

it too erases under matching with the probe. The erasure of uninterpretable features of probe and goal is the operation we called <Agree>.

BY DEFINITION, AND NOTE: THE ERASURE IS THE AGREEMENT, NOT THE MATCHING. INDEED, YOU CAN THINK OF THE MATCHING AS THE ERASURE ITSELF. THIS IMMEDIATELY MEANS YOU MUST SEPARATE AGREE IN THIS TECHNICAL SENSE FROM THE OLD CONCORD, OR OTHER FORMS OF LOOSE AGREEMENT (FOR INSTANCE OF THE SORT SEEN IN ANTECEDENCE), WHERE IT IS JUST NOT TRUE THAT MATCHING IS ERASURE, AND OF COURSE NO MOVEMENT IS INVOLVED IN ANY SENSE.

But EPP of T must also be satisfied;

OKAY, SO HERE'S WHAT WILL BE BEHIND MOVEMENT IN THE INSTANCES WHERE IT HAPPENS. NOTE, UP TO THIS POINT YOU JUST HAD AGREEMENT, THAT IS MATCHING. SO WHY IS THERE DISPLACEMENT?

in this case, by "pied-piping" of a phrase P(G) determined by the goal of T's probe, which merges with (39), becoming SPEC-T.

SO AGAIN, PIED-PIPING IS STILL ALIVE--IT'S JUST NOT 'FOR PF REASONS', AS IN CHAPTER 4. INDEED ALL THAT DISCUSSION IN CHAPTER 4 ABOUT WHAT DETERMINES THE MINIMAL PHRASE THAT PIED-PIPES WITH THE GOAL IS, SO FAR AS I UNDERSTAND, COMPLETELY RELEVANT (WHICH MEANS THE ACCOUNT IS REALLY NOT THAT DIFFERENT FROM THE ONE INVOLVING ATTRACTED FEATURES, AFTER ALL).

The combination of selection of P(G), Merge of P(G), and feature-deletion under match (Agree) is the composite operation Move, which dislocates "an unpopular candidate," eliminating all uninterpretable features.

RIGHT, JUST AS IN CHAPTER 4 IT WAS ATTRACTING OF F WITHIN SOME ELEMENT (HERE P(G)), ALSO MERGE OF P(G) (WHERE P(G) IS DETERMINED SOMEHOW OPTIMALLY), AND THE CHECKING OF AN ATTRACTING FEATURE IN THE TARGET OF MOVEMENT. DON'T GET ME WRONG: THERE ARE DIFFERENCES. FOR INSTANCE, CHAPTER 4 HAD A FEATURE CHAIN, WHEREAS THIS ARTICLE DOES NOT, PROBABLY WITH GOOD CAUSE, SINCE CERTAINLY WE HAVE INSTANCES WHERE WE DO NOT PRONOUNCE THE UPPER FEATURE (COMPLETELY CONTRARY TO THE CHAIN ASSUMPTION, SAY IN NUNES'S TERMS) AND POSSIBLY ALSO INSTANCES WHERE WE PRONOUNCE BOTH (WHICH I DISCUSSED ABOVE), BUT NOT, AT LEAST NOT THAT I KNOW, INSTANCES WHERE YOU PRONOUNCE THE UPPER FEATURE BUT NOT THE LOWER ONE, WHICH IS WHAT YOU WOULD EXPECT IF STANDARD CHAINS WERE INVOLVED.

Matching is a relation that holds of a probe P and a goal G. Not every matching pair induces Agree. To do so, G must (at least) be in the <domain> D(P) of P and satisfy locality conditions.

REMEMBER, THE NOTION OF 'DOMAIN OF P' IS WHAT USED TO BE THE 'COMPLEMENT DOMAIN' IN CHAPTER 3. LOCALITY CONDITIONS ARE WHATEVER

THOSE ARE, SAY MLC AND PERHAPS OTHERS.

The simplest assumptions for the probe-goal system are (40):

- (40)
- (I) matching is feature identity
 - (II) D(P) is the sister of P
 - (III) locality reduces to "closest c-command"

I SEE THAT THESE ARE THE ASSUMPTIONS THAT *WORK*, BUT ARE THEY ALSO THE SIMPLEST, TO SATISFY MINIMALIST SCRUPLES? (I) IS ARGUABLY SIMPLEST, AS IS (II). (III) IS TRICKY, THOUGH: THE C-COMMAND DOMAIN ALSO INCLUDES SITUATIONS WHERE THE GOAL IS INSIDE, SAY, A SUBJECT OR AN ADJUNCT TO SOMETHING IN D(P)--BUT THOSE NEVER WORK (CED EFFECTS). I ACTUALLY THINK THERE IS A MORE NATURAL NOTION THAN C-COMMAND: NAMELY 'SAME DERIVATIONAL WORKSPACE'. WE SAID THIS BEFORE: ADJUNCTS AND SUBJECTS INVOLVE SEPARATE DERIVATIONAL WORKSPACES; IF SO, IT WOULD NOT BE UNTHINKABLE THAT THE SYSTEM SHOULD BE SENSITIVE TO CLOSENESS WITHIN THE SAME DERIVATIONAL SPACE, WHICH IS LIKE SAYING THAT TWO RACERS COMPETE IF THEY ARE RACING IN THE SAME STADIUM, BUT NOT IF THEY ARE IN DIFFERENT STADIUMS--EVEN IF THOSE STADIUMS END UP JOINING THEIR TRACKS CLOSE TO THE FINISH LINE, SAY. AT ANY RATE, THE GENERAL REASONING CHOMSKY IS TRYING TO MAKE WORKS, I THINK, IF WE NARROW THINGS DOWN FROM THE UNMOTIVATED C-COMMAND (WE SAW BEFORE THAT THE COMPOSITE OF SISTER AND THE TRANSITIVE CLOSURE OF CONTAIN IS NOT A PARTICULARLY ELEGANT NOTION) TO THE MORE NATURAL 'SAME DERIVATIONAL SPACE.' I'LL USE THE TERM 'DERIVATIONAL CASCADE' TO REFER TO WHAT GOES ON WITHIN A GIVEN DERIVATIONAL WORKSPACE, PARTICULARLY AS IT BLEEDS INTO THE INTERPRETIVE COMPONENTS.

Thus D(P) is the c-command domain of P, and a matching feature G is <closest to> P if there is no G' in D(P) matching P s.t. G is in D(G').

THE SAME HOLDS FOR THE NARROWER, MORE MOTIVATED VARIANT I'VE SUGGESTED.

In the absence of evidence to the contrary, we adopt (40), with a qualification taken over from earlier work:

{See <MP>, Ura (1996). The condition stated there refers only to SPECs of the same head. Whether the generalization (41) is appropriate depends on answers to questions about the structure of more complex constructions (double-object verbs, etc.). These and other questions, including parametrization, are put aside here. See Boskovic (1997), McGinnis (1998), among others.

ANOTHER PAPER TO SEE IN THIS RESPECT IS ORMAZABAL'S (1999).

The notion "feature occurrence" used implicitly here for expository convenience can be eliminated by restatement in terms of the heads to which features belong. }

TRUE, BUT A DEEP ISSUE REMAINS: ARE THERE ANY OCCURRENCES OF FEATURES, AND IF NOT WHY NOT? ANOTHER WAY TO SAY THIS IS WHY ARE FEATURES DIFFERENT FROM CATEGORIES, THE LATTER CLEARLY ENTERING INTO VARIOUS OCCURRENCES AS THE DERIVATIONS PROCEED? YOU CAN OF COURSE STIPULATE THAT DERIVATIONS ONLY MANIPULATE CATEGORIES, BUT YOU WANT TO UNDERSTAND WHY THINGS ARE THAT WAY, WHAT IS NECESSARY ABOUT THAT. FOR INSTANCE, COULD IT BE THAT THE INTENTIONAL SYSTEM CAN ONLY MANIPULATE CATEGORIES, NOT MERE FEATURES FROM THE CONCEPTUAL SYSTEM? I'M NOT SAYING THAT'S THE ANSWER, BUT RATHER THAT YOU'D NEED AN ANSWER OF THAT SORT.

(41) Terms of the same minimal domain are "equidistant" to probes

NOTICE, ALSO, THAT JUST AS WE HAD NO REASON FOR THIS IN CHAPTER 3, WE STILL DON'T HAVE IT IN THIS ARTICLE. THIS IS A GENERALIZATION THAT MUST BE UNDERSTOOD TO FOLLOW FROM SOMETHING ELSE, NOT ANYTHING OBVIOUS. IT SEEMS AS IF FOR THE PURPOSES OF THE DERIVATION DISTANCE COLLAPSES WITHIN A GIVEN CHUNK, AS IF THE INNER STRUCTURE OF THIS CHUNK WAS IRRELEVANT. OBVIOUSLY, THAT'S ANOTHER POCKET OF REGULARITY WHERE YOU EXPECT CONSERVATIONS TO TAKE PLACE (IN THIS INSTANCE, PERHAPS L-RELATEDNESS).

The minimal domain of a head H is the set of terms immediately contained in projections of H.

AS BEFORE, THIS IS A SET-THEORETIC OBJECT, WHICH CONFIRMS THE IDEA THAT THIS IS A NEW DOMAIN OF CONSERVATION, SEPARATE FROM CONFIGURATIONAL MATTERS--THAT IS, A HIGHER ORDER OBJECT BUILT FROM A PHRASAL OBJECT. SO FAR, WE HAVE SEEN THE FOLLOWING DOMAINS OF CONSERVATION: WORDS, PHRASES, MINIMAL DOMAINS, AND POSSIBLY ALSO NUMERATIONS (MORE ON THIS BELOW, WHERE THE ISSUE BECOMES CLEARER). TO SERIOUSLY IMPLEMENT THE IDEA THAT CHAINS ARE OBJECTS IN DIFFERENT POSITIONS AT THE SAME TIME, THEN YOU ALSO NEED A FIFTH DOMAIN OF CONSERVATION, A CHAIN LEVEL ONE. I SHOULD EMPHASIZE THAT I'M NOT ADDING ANYTHING HERE, I'M JUST TRYING TO CODE IN REGULAR WAYS THE ASSUMPTIONS THAT CHOMSKY HAS BEEN MAKING. THAT IS, YOU CANNOT JUST ASSUME (40) AND GO TO THE NEXT PAGE; IN EFFECT, WHAT YOU'VE ASSUMED IS THE EXISTENCE OF A KIND OF ENTITY WHERE INNER PHRASAL STRUCTURE IS IRRELEVANT. SIMILARLY WITH ALL THE OTHER DOMAINS WE'VE BEEN ASSUMING, WHERE CERTAIN INTEGRITY HOLDS; FOR INSTANCE, WITHIN WORDS YOU DON'T CARE ABOUT LCA, SAY, AND WITHIN CHAINS YOU MUST MAKE SURE THAT ONLY ONE OF THE CHAIN SITES GETS INTERPRETED, AND SO ON. YOU CAN EITHER DO THIS MECHANICALLY WITHOUT OBSERVING WHAT'S COMMON TO

ALL THESE INSTANCES, OR YOU CAN CHOOSE TO REFLECT ON THE COMMONALITIES. THE REASON I USE THE CONSERVATION METAPHOR IS TWOFOLD. FIRST, IN PHYSICS A SIMILAR METHODOLOGY WAS USEFUL WHEN PEOPLE WERE TRYING TO SORT OUT THE BEHAVIOR OF NEW PARTICLES AND THEIR INTERACTIONS, THE RESULT OF WHICH WAS GELMAN'S QUARK MODEL. SECOND, AND MOST IMPORTANTLY, IT SEEMS AS IF WE DO WANT TO CONSERVE CERTAIN PATTERNS IN THE DOMAINS I'M ASSUMING, EVEN IF WE DO IT TACITLY. WE'VE SEEN THIS FOR THE TOKEN-TO-CHAIN MAPPING, AND IF, FOR INSTANCE, L-RELATEDNESS IS WHAT CHARACTERIZES THE MINIMAL DOMAIN ESSENCE, THEN THE FACT THAT WE PRESERVE L-RELATEDNESS THROUGHOUT THE DERIVATION WOULD BE THE OBVIOUS RAISON D'ETRE OF MINIMAL DOMAINS-- AND WITH IT WOULD POSSIBLY COME YOUR DEFINITION OF A VS. A' MOVEMENT, ANOTHER MANIFESTATION OF THE CONSERVATION OF L-RELATEDNESS. IT IS WORTH POINTING OUT ALSO THAT THIS ARCHITECTURE IS PARADIGMATIC, IN THE SENSE THAT IT ESTABLISHES VERTICAL CUTS ON THE SYSTEM, EACH OF WHICH IS PROGRESSIVELY DEFINED ON THE PREVIOUS. FOR INSTANCE, CHAINS ARE NECESSARILY DEFINED IN TERMS OF PHRASES, BUT NOT THE OTHER WAY AROUND, AND PHRASES IN TERMS OF WORDS, AND NOT VICE-VERSA. THIS DIMENSIONAL TAKE ON THE LINGUISTIC ARCHITECTURE (WHICH IS OF COURSE FULLY COMPATIBLE WITH THE HORIZONTAL, OR SINTAGMATIC CUT THAT WE USUALLY TAKE) WILL BE OF CERTAIN IMPORTANCE LATER ON IN THE SEMESTER.

With matching restricted to identity, Case and lexical category cannot enter into Agree or Move, since the probes do not manifest these features.

THAT IS, YOU DO NOT FIND NOMINATIVE LOOKING FEATURES, SAY, IN TENSE; WHAT YOU TYPICALLY FIND INSTEAD IS THE ABILITY TO RELATE TO NOMINATIVE FEATURES IN A NOMINAL. AND THE SAME WITH PREPOSITIONS AND, AS FAR AS I KNOW, THE REST--FOR CASE. SIMILAR CONSIDERATIONS OBTAIN FOR CATEGORIAL FEATURES. THEN IT FOLLOWS THAT WHAT GEARS MOVEMENT/AGREE IS NOT CASE OR CATEGORIAL DEPENDENCY, BUT JUST AGREEMENT. THE SAME IS TRUE FOR OBJECTS:

And Object Shift must be for (here invisible) object agreement, with ancillary Case checking.

If uninterpretable features serve to implement operations, we expect that it is structural Case that enables the closest goal G to select P(G) to satisfy EPP by Merge.

THAT IS, SOME GOAL G ENTERS INTO AGREEMENT WITH SOME PROBE P IN A CATEGORY K (WHEREBY P(=G) ERASES); THEN THE IDEA IS THAT ALL CASE DOES IS SELECT OR (SINCE, AGAIN, I WOULDN'T USE THAT TERM, I'LL SAY:;) PICK OUT A PHRASE P(G) TO SATISFY EPP IN THE DOMAIN OF K BY COPY+MERGE. CASE IS JUST THE FRAME THAT TELLS YOU WHAT WILL SATISFY EPP IN THIS LANGUAGE--HOW, WE DON'T KNOW YET.

Thus, if structural Case has already been checked (deleted), the phrase P(G) is "frozen in place," unable to move further to satisfy EPP in a higher position.

MECHANICALLY, THIS WILL BE IMPORTANT, PREDICTING A NUMBER OF INSTANCES WHERE NPS ARE STUCK AFTER THEY CHECK CASE, ERASING THE CORRESPONDING CASE FEATURE (I WOULD HAVE USED THE WORD 'ERASE' HERE TOO, INSTEAD OF THE CONFUSING TERM 'DELETE'). HERE'S ALSO A MECHANICALLY CONFUSING ISSUE. WHEN P AND G ENTER INTO A CHECKING RELATION, WHERE P=G, IT IS P THAT ERASES (SUICIDAL GREED). WHAT ABOUT CASE THEN? WELL, CASE IS A FEATURE OF P(G), A PHRASE CONTAINING G--AND IT TOO ERASES. THE ONLY FEATURE THAT REMAINS IS G. YOU CAN THINK OF THE FREEZING EFFECT AS FOLLOWS: SINCE CASE IS GONE FROM P(G), IF I TRIED TO MATCH G (WHICH IS NOT ERASED, REMEMBER) TO SOME HIGHER P' LATER ON, I WOULDN'T BE ABLE TO SATISFY EPP IN TERMS OF MOVING THE PHRASE P(G), SINCE I HAVE NO CASE FEATURE TO MAKE IT ACCESSIBLE TO THE SYSTEM.

More generally, uninterpretable features render the goal <active>, able to implement an operation: to select a phrase for Merge (pied-piping) or to delete the probe.

THIS IS GOING TO BE THE GENERAL LINE FOR *ALL* UNINTERPRETABLE FEATURES (FOR INSTANCE, A SIMILAR MECHANISM WILL BE POSTULATED FOR WH-MOVEMENT). AN ACTIVE GOAL IS, FIRST, ONE THAT LETS YOU PICK OUT A PHRASE FOR EPP. BUT CHOMSKY SEEMS TO BE SUGGESTING HERE, ALSO, THAT IT MAY BE ALLOWING YOU TO DELETE THE PROBE AS WELL; THAT CLEARLY WOULD NOT BE THE CASE OF CASE, SINCE YOU'LL NEVER FIND AN IDENTICAL CASE FEATURE UPSTAIRS--BUT COULD IT BE THE CASE OF A WH-FEATURE? PERHAPS, IT DEPENDS ON HOW YOU INTERPRET WH-FEATURES, AS INTERPRETABLE (AS IN CHAPTER 4) OR UNINTERPRETABLE (AS HERE).

The operations Agree and Move require a goal that is both local and active.

LOCALITY WE'VE ASSUMED ALL ALONG, BUT NOW WE'RE ADDING ACTIVITY. I SHOULD SAY, PERHAPS, THAT WHILE THIS INTERPRETATION OF CASE IS ENTIRELY PLAUSIBLE, IT IS BY NO MEANS OBVIOUSLY NECESSARY, UNLESS WE MAKE SOME NON-TRIVIAL ASSUMPTIONS ABOUT PHRASES. LET'S THINK ABOUT THIS FOR A SECOND. PHRASE-MARKERS, YOU RECALL, ARE OBJECTS OF THE FORM {A, {B, C}}, WHERE TERMS B AND C CAN BE ARBITRARILY COMPLEX, AND LABEL A CAN BE THOUGHT OF AS A GENERALIZATION, NOTHING OBVIOUSLY REAL TO THE SYSTEM--CERTAINLY NOT ANYTHING ACCESSIBLE TO IT. IF SO, IMAGINE A PHRASE-MARKER WHERE C CONTAINS A COMPLEX SET OF TERMS. HOW DOES THE SYSTEM KNOW THAT IT IS SUPPOSED TO WORK WITH A CERTAIN CRUCIAL TERM P(G)? FOR THE PROBE-GOAL RELATION THE ISSUE IS NOT SO TROUBLING IF P JUST SEEKS SOMETHING LIKE IT ITSELF. BUT WHAT IS SOUGHT FOR EPP FEATURES? DOES THE SYSTEM HAVE THE ABILITY TO LOOK FOR SOMETHING LIKE 'THE MINIMAL VALID PHRASE THAT CONTAINS G?' PERHAPS IT

DOES, ALTHOUGH THAT'S A HIGHLY ARTICULATED REASONING. PERHAPS, INSTEAD, THE SYSTEM JUST SEEKS 'A STUPID UNINTERPRETABLE CODING'. THAT WOULD BE CASE. SO IT ALL DEPENDS ON WHAT KIND OF INTELLIGENCE YOU'RE WILLING TO GIVE THE SYSTEM. (AS USUAL, AND AS IT SHOULD BE FOR THIS KIND OF NATURAL SYSTEM, WE'RE GOING WITH THE DUMBEST PROCESS THAT DOES THE TRICK.)

{Among the problems that arise is the status of scrambling. The logic would suggest that for at least some cases, a scrambling feature induces pied-piping even after Case assignment, with the pied-piped element "attracted" by a higher probe,

THAT IS, GIVEN THAT SOME INSTANCES OF SCRAMBLING SEEM TO INVOLVE MOVEMENT AFTER CASE HAS ALREADY BEEN CHECKED, IF THEY ARE TO REDUCE TO THE PRESENT SYSTEM, ONE WOULD HAVE TO POSTULATE SOME KIND OF SCRAMBLING FEATURE THAT DOES THE JOB OF CASE. THIS IS SOMEWHAT DUBIOUS, IN THAT, SO FAR AS I KNOW, THE RELEVANT SCRAMBLING FEATURE HAS NEVER BEEN ATTESTED--AND IS PROPOSED HERE ONLY IN TERMS OF THE LOGIC.

while other cases fall into a category distinct from feature-driven movement.

THAT IS INDEED MORE PLAUSIBLE, SUGGESTING SCRAMBLING FALLS TOGETHER WITH OTHER INFORMATION-DRIVEN, NON-INTENTIONAL OPERATIONS.

For exploration of alternatives in a comparative study, see Sauerland (1998a).}

We therefore have the possibility of <defective intervention constraints> in a structure (42), where > is c-command, β and $_$ match the probe $_$, but β is inactive so that the effects of matching are blocked:

(42) $_ > \beta > _$

THAT IS, THE FACT THAT β IS INACTIVE BECAUSE ITS CASE HAS BEEN CHECKED RAISES THE QUESTION: DO WE THEN HAVE A RELATION BETWEEN $_$ AND $_$ WHICH WOULD OTHERWISE BE BLOCKED? CHOMSKY THINKS WE DO. A RELATED EXAMPLE TO THINK ABOUT WOULD BE THIS: IMAGINE β PROBES INTO $_$, AND FINDS IT AS ITS GOAL, WITH β THEN DELETING AND P($_$) DOING ITS BUSINESS EPP-WISE, THUS ELIMINATING THE CORRESPONDING CASE FEATURE IN $_$; DOES THAT MEAN, THEN, THAT $_$ COULD NOT BE A VALID GOAL FOR A HIGHER $_$, EVEN IF $_$ SATISFIES ITS EPP FEATURE DIFFERENTLY? IF I READ HIM CORRECTLY, CHOMSKY THINKS NOT. IF SO, THIS WOULD BE RATHER STRONG EVIDENCE FOR THE IDEA THAT ACTIVITY IS AN ISSUE IN ALLOWING LONG-DISTANCE OPERATIONS. THAT *WOULD* BE RATHER DIFFERENT FROM THE R&U THEORY, SINCE THERE CASE WAS NOT WHAT SANCTIONED THE AGREEMENT RELATION, BUT RATHER SOMETHING THAT WAS ASSIGNED IN THE RELEVANT CONFIGURATIONS. IN ANY CASE:

We return to some illustrations. {For a similar configuration in phonology, with > = linear order, see the discussion in Halle (1995) of coronal assimilation in Sanskrit, barred by intervening (non-assimilating) coronal. }

In <MP>, Agree is analyzed in terms of feature-movement (Attract) and a concept of matching that is left unclear. Here we take matching to be identity and dispense with Attract, with complications it induces about extended MLIs, feature chains, and other matters. Checking reduces to deletion under matching with an active local goal and ancillary deletion of the uninterpretable feature that rendered the goal active. I will use the terms "checking" and "attract" only for convenience.

THIS IS PRETTY MUCH A SUMMARY OF WHAT WE'VE DONE SO FAR, AND ONE OF THE MAIN CHANGES FROM CHAPTER 4.

Suppose that the EPP feature of T could be satisfied more simply than by the full operation Move.

HERE WE GO INTO EXPLETIVES, AND NOTE THE FAST, THOUGH ACCURATE ASSUMPTION: MERGE IS SIMPLER THAN MOVE, SO IF YOU COULD SATISFY EPP BY MERGE, THAT DERIVATION WINS OVER ITS ALTERNATIVE, WHICH HAS TO INVOLVE MERGE PLUS COPY (AND POSSIBLY DELETION, DEPENDING ON HOW YOU SEE CHAINS).

That is the case in (35ii), repeated here:

(43) there [_{<_>} T-was elected an unpopular candidate]

Here the lexical array includes the expletive <there>.

THAT IS A CRUCIAL ASSUMPTION, OR ELSE YOU WOULD NEVER GET an unpopular candidate was elected, WHICH IS POSSIBLE IN SPITE OF INVOLVING MOVEMENT BECAUSE IT DOES NOT COMPETE WITH A DERIVATION INVOLVING AN EXPLETIVE.

At stage _{<_>} of the derivation (= (39)), the independent operations Agree and pure Merge suffice: Agree deletes the _{<_>}-set of T and the structural Case of <candidate>, and Merge (of <there>) satisfies the EPP feature of T. The more complex operation Move is preempted; dislocation does not take place, though we have long-distance agreement of T and its goal (its associate).

{On the assumptions of <MP> 4.10, multiple-subject (including transitive expletive) options are parametrized in terms of deletion of EPP-features. See particularly Ura (1996), and for skepticism about the option, Zwart (1997).

IN ANY CASE, THE ISSUE IS REAL, HOWEVER YOU CODE IT. EXPLETIVES IN ENGLISH APPEAR ONLY (BASICALLY) WITH UNACCUSATIVES AND PASSIVES,

WHEREAS THEY GO WITH TRANSITIVES IN MANY OTHER LANGUAGES. URA'S POINT WAS THAT THIS CORRELATES WITH MULTIPLE SUBJECTS OF THE SORT SEEN IN EAST ASIAN LANGUAGES. THE EXTENSION TO GERMANIC LANGUAGES, HOWEVER, IS FAR FROM OBVIOUS (IT DOESN'T SEEM AS IF THESE HAVE BONA FIDE MULTIPLE SUBJECTS, PUTTING ASIDE EXPLETIVES).

Agreement in the sense discussed here is to be distinguished from concord, with different properties. }

WE ALREADY SAID THIS BEFORE. CONCORD, FOR INSTANCE, IS WHAT GOES ON BETWEEN A NOUN AND ITS ADJECTIVES--THOUGH THE ISSUES ARE VERY MURKY.

Manifestation of structural Case depends on interpretable features of the probe: finite T (nominative), <v> (accusative), control T (null), on our earlier assumptions.

AGAIN, THIS SEEMS TO BE TRUE, BUT WOULDN'T FOLLOW FROM ANYTHING SAID SO FAR. THAT IS, THERE IS NOTHING LOGICALLY INCONCEIVABLE ABOUT FL MARKING ALL GOALS IN TERMS OF A SINGLE, DESIGNATED CASE.

We may therefore regard structural Case as a single undifferentiated feature.

INDEED, THIS WILL BE CRUCIAL ON EMPIRICAL GROUNDS, GIVEN THAT WE WANT DIFFERENT CASES TO COUNT FOR INTERVENTION PURPOSES.

The same would be expected for the uninterpretable $_set$ of the probe. Its manifestation depends on interpretable features (namely, $_features$) of the goal, so that it too can be taken to be undifferentiated as to the value of the individual features of the $_set$ ([+/-plural], etc.).

UNFORTUNATELY, THE COMPARISON IS NOT TOTALLY FAIR. THAT IS, THE FACT THAT SOMETHING MANIFESTS ITSELF AS PLURAL OR MASCULINE DOES SEEM TO DEPEND ON GENUINE INTERPRETIVE FEATURES OF THE CATEGORY THAT BEARS THOSE FEATURES, AS MANIFESTED IN THE SEMANTICS. CASE IS VERY DIFFERENT. THERE IS NO SUCH THING AS A SEMANTIC NEED FOR T TO BEAR NOMINATIVE CASE. SO A MYSTERY REMAINS: WHY WE HAVE DIFFERENT CASE *VALUES*, EVEN IF WE ACCEPT THE NEED FOR CASE FEATURES IN TERMS OF ACTIVITY.

For both probe and goal, the form of the uninterpretable features is determined by Agree. To rephrase in traditional terms, verbs agree with nouns, not conversely, and Case is assigned.

THE TRADITIONAL OBSERVATION SEEMS CORRECT, AND PRESENT CONJECTURES CERTAINLY ARE ON TRACK WITH REGARDS TO AGREEMENT, BUT THE FACT THAT 'CASE IS ASSIGNED', WHILE TRUE DOESN'T SEEM TO FOLLOW.

We therefore understand "feature identity" in (I) of (40) to be identity of the choice of feature,

not of value.

OF COURSE, THAT'S CRUCIAL, AND WHAT WE'RE SAYING. ANOTHER WAY OF PUTTING THIS IS AS FOLLOWS: THERE IS NO SUCH THING AS AN ACCUSATIVE FEATURE, ANYMORE THAN THERE IS A PLUS FEATURE. ACCUSATIVE IS A VALUE OF A CASE FEATURE, JUST LIKE PLUS IS A VALUE OF A NUMBER FEATURE. FEATURE IDENTITY, THUS, IS NOT SENSITIVE TO HOW THE FEATURE MANIFESTS ITSELF. THIS IS LIKE SAYING THAT THE COMPUTATIONAL SYSTEM DOESN'T MANIPULATE SUBSTANTIVE INSTANCIATIONS OF THE CONCEPTUAL SYSTEM, BUT MERELY ITS GENERAL PARAMETERS. THIS IS VERY PLAUSIBLE. JUST AS THE GRAMMAR DOESN'T MANIPULATE VALUES OF SYNTACTIC PARAMETERS (THERE IS NO PRINCIPLE OF GRAMMAR THAT DEPENDS ON WHETHER YOU ARE OR ARE NOT PRO-DROP), SO TOO THE GRAMMAR DOESN'T MANIPULATE VALUES OF CONCEPTUAL PARAMETERS INVOLVED IN LEXICAL CHARACTERIZATION. SYNTAX IS PURELY FORMAL, AND HAS NO WAY OF KNOWING, FOR INSTANCE, WHETHER IT IS OPERATING WITH *he* OR *she* IN *she loves peanuts* (OR FOR THAT MATTER, WITH *loves* AS OPPOSED TO *hates*, AND SO ON). IF THIS IS GENERALLY THE CASE, WE MIGHT HAVE A HINT HERE AS TO WHY WE HAVE CASE *VALUES* TO BEGIN WITH, FOR IN A SENTENCE LIKE *he loves her* THE GRAMMAR JUST SEES SOMETHING ROUGHLY LIKE [D [v [V D]]]. INDEED, IN THE NUMERATION YOU'D HAVE TWO D TOKENS, BUT HOW DO YOU CODE THAT? PERHAPS CASE VALUATION IS NOTHING BUT A WAY OF CODING THE D TOKENS, SO THAT IN THE NUMERATION YOU HAVE {[D-nom], [D-acc], V, v,...}, AND THEN IN THE SYNTAX YOU OPERATE WITH [D-nom [v [V D-acc]]]. THAT WOULD JUSTIFY WHY YOU HAVE DIFFERENT VALUES. OF COURSE, THAT STILL DOESN'T JUSTIFY WHY YOU HAVE THE ACCUSATIVE VALUE ASSOCIATED TO *v* AND THE NOMINATIVE VALUE ASSOCIATED TO *T*, AND SO ON. YOU MIGHT THINK THIS IS JUST NOMENCLATURE, BUT IT ISN'T. LANGUAGES DIFFER IN WHICH OF THESE THEY TAKE TO BE THE DEFAULT VALUE (E.G. IN TERMS OF CITATION FORMS OR TOPIC/DEFAULT CASE), AND ALSO IN HOW THEY DISTRIBUTE THESE CASE VALUES TO DIFFERENT ARGUMENTS. WE SHOULD RETURN TO THESE ISSUES.

More important, defective intervention effects are induced whether or not β and $_$ of (42) are identical in $_$ -feature value (singular blocks plural agreement, etc.).

THIS IS WHAT WE SAID ABOVE, AND ONE OF THE MAIN EMPIRICAL ARGUMENTS NOT TO CONSIDER A VALUED FEATURE WHAT IS RELEVANT IN SYNTAX; MORE GENERALLY (A POINT THAT GOES BACK TO BRESNAN), VALUED FEATURES NEVER DO ANYTHING FOR YOU: NO TRANSFORMATION OR ANY OTHER RULE IS GEARED TOWARDS VALUED FEATURES.

This lends theory-internal support to the earlier observation that $_$ -features are interpretable only for N; their value is specified only in this case.

THIS IS A VERY INDIRECT ARGUMENT, SINCE UP TO NOW WE WERE TALKING ABOUT THE VISIBILITY OF FEATURE VALUES FOR THE SYNTAX IN GENERAL. THE

POINT IS: IN VERBS, YOU DON'T CARE ABOUT AGREEMENT VALUES, SINCE YOU GET INTERVENTION EFFECTS REGARDLESS OF VALUES. THEREFORE (?) INTERPRETABILITY OF VALUES IS ONLY IN N. THE THEREFORE, HOWEVER, DOESN'T LOGICALLY FOLLOW (EVEN IF IT SEEMS TRUE).

Notice also that only the most underspecified element, PRO, can have null Case,

THIS SOUNDS PROFOUND, BUT IT ISN'T. IT IS ARGUABLE THAT PRO IS THE MOST UNDERSPECIFIED ELEMENT, IF IT INDEED EXISTS, BUT IT ISN'T OBVIOUS EITHER THAT NULL CASE IS IN ANY SENSE UNDERSPECIFIED (IT'S JUST NOT PRONOUNCED, IF IT INDEED EXISTS), OR THAT THERE SHOULD BE A CORRELATION IN SPECIFICATION HERE. IF THAT CORRELATION HOLDS, IT HAS TO BE ASSUMED FROM SOMETHING ELSE, NOT SAID EXPLICITLY HERE. WE SHOULD RETURN TO THIS TOO.

so raising of \bar{y} PRO to SPEC-T crashes the derivation when T is a control infinitival.

OF COURSE, THAT'S WHAT WE WANT TO ACHIEVE, BUT LET'S NOT CONFUSE THE DESIRE WITH AN EXPLANATION.

We take deletion to be a "one fell swoop" operation, dealing with the \bar{y} -set as a unit.

AGAIN, THIS DOESN'T FOLLOW FROM ANYTHING OBVIOUS, ALTHOUGH IT SEEMS TRUE. YOU CAN BUILD THE GRAMMAR THAT WAY, OF COURSE, BUT YOU HAVE TO WONDER WHETHER IT IS NECESSARY.

Its features cannot selectively delete: either all or none. The \bar{y} -features of T do not agree with different NPs, for example.

LET'S SEE IF WE CAN ACTUALLY BUILD A RELEVANT EXAMPLE. IMAGINE A LANGUAGE THAT HAS SUBJECT AND OBJECT AGREEMENT, WHERE YOU CAN ACTUALLY TEST THESE ISSUES. AND IMAGINE A TENSE THAT HAS AGREEMENT IN PERSON WITH THE SUBJECT AND AGREEMENT IN NUMBER WITH THE OBJECT. FOR EXAMPLE, IN BASQUE YOU CAN SAY *Nik Miren maite dut*, LITERALLY I-subject Miren-object love I/SG Aux III/SG. NOW IMAGINE THE FOLLOWING SENTENCE: *Nik Miren maite naiz*, THAT IS, LITERALLY AGAIN I-subject Miren-object love I/SG Aux. WHY COULDN'T I USE THIS FIRST PERSON SINGULAR AGREEMENT TO AGREE IN FIRST PERSON WITH THE SUBJECT AND IN NUMBER WITH THE OBJECT? IF I COULD DO THAT, I'D PROCEED TO CHECK THE SUBJECT AND OBJECT CASE FEATURES WITHOUT ANY PROBLEM. BUT THE SENTENCE IS CRASHINGLY UNGRAMMATICAL. IF NONE OF YOU FINDS A LOOPHOLE WITH MY ARGUMENT (AND NOTE THAT PRESUMABLY I AND Miren ARE BOTH WITHIN THE SAME MINIMAL DOMAIN), THEN THIS IS THE SORT OF EXAMPLE THAT PROVES CHOMSKY'S POINT.

In the same spirit, we assume that only a probe with a full complement of $_features$ is capable of deleting the feature that activates the matched goal.

THIS ONE IS EVEN MORE STIPULATION THAN THE PREVIOUS, EVEN IF RIGHT. WE NEED TO SAY IT, IN ORDER NOT TO ELIMINATE CASE FEATURES IN PLACES WHERE WE WANT THEM TO REMAIN ACTIVE; FOR INSTANCE:

Suppose that the probe for participial (like adjectival) $_$ is a $_set$ lacking the feature [person], and that G is the closest matching goal in its search space: $P(G) = DP$ may be attracted to SPEC- $_$ deleting the probe of $_$ (participial agreement), but the operation will not delete structural Case in DP, which can move on to SPEC-T, deleting the probe of T and Case of DP (subject agreement).

LET'S ILLUSTRATE. TAKE THE ROMANCE *ellos han sido arrestados*, LITERALLY they have-III/pl. been arrested-pl. YOU WANT TO BE ABLE TO SAY THAT they MOVES TO A SITE WHERE IT CHECKS NUMBER AGREEMENT WITH THE PARTICIPIAL (WHICH THEN ERASES), BUT STILL CAN MOVE ON TO T IN ORDER TO CHECK THE FULL SET OF $_FEATURES$ (ALSO INCLUDING PERSON). THE LOGIC THEN DICTATES THAT THE CASE FEATURE IS STILL ACTIVATING THE SUBJECT, OR IT WOULDN'T MOVE ANY FURTHER.

<v> and nondefective T, with a full complement of $_features$, delete the uninterpretable feature that activates the matched goal (raised or not).

THAT'S THE BOTTOM LINE, BUT AGAIN NOTHING PROFOUND SEEMS TO GRANT THIS. IT IS CERTAINLY THE CASE THAT, EMPIRICALLY, THOSE TWO (AND ALSO DATIVE SITES, WHICH I'M SURE WE COULD ADAPT TO THE PRESENT SYSTEM) IS WHERE PERSON FEATURES ARE ENCODED. IT THUS SEEMS AS IF THE PERSON-ENCODING SITES ARE CENTRAL IN ELIMINATING THE CASE FEATURES (INDEED, ORMAZABAL (1999) ARGUES THAT WHAT MATTERS IN ALL THESE SORTS OF INSTANCES IS NOT CASE, BUT PERSON INSTEAD, WHICH SEEMS LIKE A GOOD WAY OF BITING THE BULLET--EVEN IF THE MYSTERY IS STILL THERE: WHY DO THOSE SITES CODE PERSON AND WHY DO WE BOTHER TO CODE PERSON DISTINCTIONS?). IT IS PERHAPS WORTH MENTIONING, ALSO, THAT THE PERSON FEATURE SEEMS CRUCIAL IN 'LICENING' *pro*, AS ARGUED SOME TIME AGO BY GALVES, FOR INSTANCE. AS YOU PROBABLY KNOW, *pro* TYPICALLY MANIFESTS ITSELF IN CASE SITES, WHICH SUGGESTS A TIGHT CORRELATION PERSON/CASE/*pro*. WHETHER IT IS IN TERMS OF FULL SETS OF $_FEATURES$, AS IN CHOMSKY'S SPECULATION, I DON'T KNOW.

{The analysis of structural Case is along the lines of George and Kornfilt (1981). As they observe, structural Case linked to $_features$ may be dissociated from finiteness.

A SIMILAR POINT WAS RAISED IN RAPOSO'S THESIS CONSIDERABLY BEFORE THAT (1973), AND AGAIN IN HIS LI ARTICLE ON INFLECTED INFINITIVALS.

Matters become more complex when we consider ergative/absolute and mixed systems, and

languages in which $\bar{_}$ -features without finiteness do not suffice for nominative Case assignment (see Iatridou 1993).}

How would non-control infinitivals (T<def>) and weak expletives EXPL of the <there>-type fit into this picture? The former category falls into place if T always has at least a minimal feature complement, perhaps only [person] for T<def>. If so, Move of $\bar{_}$ to SPEC-T<def> will delete the $\bar{_}$ -set of T (= uninterpretable [person]) but not the structural Case feature of $\bar{_}$, so that $\bar{_}$ can undergo further movement and agreement.

GIVEN WHAT I JUST SAID ABOUT PERSON, I FIND IT HARD TO BELIEVE THAT DEFECTIVE T HAS PRECISELY THE FEATURE THAT SEEMS SO INVOLVED IN THE CASE SYSTEM. INDEED, IN LANGUAGES WITH INFLECTED INFINITIVALS, SUCH AS PORTUGUESE, WHAT YOU SEEM TO BE ADDING TO THESE INSTANCES WITH DEFECTIVE T IS PRECISELY A PERSON FEATURE, AND THEN ALL OF A SUDDEN CASE LICENSING BECOMES POSSIBLE AND FURTHER RAISING IS IMPOSSIBLE. THIS IS JUST AS CHOMSKY WOULD WANT IT, BUT ONLY IF HE DROPS THE IN-PASSING ASSUMPTION THAT THE DEFECTIVE FEATURE IN QUESTION IS PERSON. OF COURSE, HE NEEDS A FEATURE IF HE WANTS TO MOTIVATE SUCCESSIVE-CYCLIC MOVEMENT, AND GIVEN THE LOGIC OF WHAT HE'S SAID SO FAR, THE FEATURE IN DEFECTIVE T CANNOT BE, OBVIOUSLY, A CASE FEATURE OR (LESS OBVIOUSLY, BUT WITH THE SAME REASONING) A CATEGORIAL FEATURE OF THE SORT HE INVOLVED IN CHAPTER 4. THERE IS NOT MUCH MORE LEFT: NUMBER AND GENDER FEATURES DON'T SEEM TO BE INVOLVED HERE, CONTRARY TO WHAT HAPPENS IN PARTICIPIALS, WHERE YOU FIND PRECISELY THOSE. SO EITHER THERE IS A NEW, MYSTERIOUS FEATURE IN THESE INSTANCES, OR THERE ISN'T SUCCESSIVE CYCLICITY HERE, OR ELSE SUCCESSIVE CYCLICITY IS AN ENTIRELY DIFFERENT PROCESS. WE'LL RETURN TO ALL OF THIS.

The phase heads <v>/C have no counterpart to T<def> with a reduced $\bar{_}$ -set, and therefore do not provide an "escape hatch" for successive-cyclic A-movement.

THE ASSUMPTIONS HERE ARE SEVERAL. FIRST THAT SUCCESSIVE CYCLIC MOVEMENT IS DONE THROUGH THE SPEC OF v , C, OR T (THIS MIGHT SEEM OBVIOUS, BUT AGAIN IS NOT NECESSARY). SECOND, THAT ONLY DEFECTIVE T HAS A REDUCED $\bar{_}$ -SET (THAT, WE CAN GRANT). THIRD, THAT MOVEMENT OF TYPE X ACROSS A REDUCED X-SET DOESN'T RESULT IN THE DE-ACTIVATION OF THE RELEVANT PHRASE (THIS IS THE NEW IDEA, AN INTERESTING ONE, BUT ONE THAT HAS TO BE DERIVED WITHIN THE SYSTEM). IF YOU MAKE THESE ASSUMPTIONS, THE NORMAL ESCAPE HATCH FOR A-MOVEMENT WILL BE THE DEFECTIVE T (ON SIMILAR GROUNDS, v AND C WILL BE THE ESCAPE HATCHES FOR A' MOVEMENT, AT LEAST THOSE v 'S AND C'S THAT HAVE DEFECTIVE A'-FEATURES, WHATEVER THOSE ARE).

Weak EXPL shares the basic movement/attraction properties of nominals. That is expected if EXPL has an uninterpretable feature F that activates it until erased and a $\bar{_}$ -set G that matches a probe in T. But G is uninterpretable for EXPL, so a distinct F is unnecessary, obviating the need for structural Case in EXPL.

SO BY THE FACT THAT THE EXPLETIVE IS AN EXPLETIVE, IT DOESN'T NEED CASE--THIS IS CLEVER, OF COURSE.

The composition of G is determined by two conditions: (A) EXPL can raise to SPEC-T<def>; (B) EXPL cannot delete the probe of nondefective T.

THESE ARE EMPIRICAL CONDITIONS. A) IS THE STANDARD ASSUMPTION, AND B IS NEEDED, I BELIEVE, IN ORDER TO INSTANTIATE THE EXPLETIVE/ASSOCIATE RELATION.

(A) requires that G contain a feature to match the probe of T<def> ([person], if what precedes is correct).

OR WHATEVER, IF WHAT PRECEDES IS NOT. IN FACT, IT SEEMS BIZARRE TO CLAIM THAT AN EXPLETIVE HAS A PERSON FEATURE.

From (B) it follows that G must be less than a full $_set$, hence optimally just [person]. That (B) holds is shown by long distance agreement structures such as (44ii-iv), (ii) surfacing commonly as (iii), or in English more naturally as (iv), as noted:

- (44)
- (i) they declared [three men guilty]
 - (ii) there were declared [three men guilty]
 - (iii) there were declared guilty three men
 - (iv) there were three men declared guilty

If the matching feature of the probe were deleted by the operation, it would not be available for associate matching and the nominative Case of the associate would remain unchecked because of the lack of a full complement of features in T (compare participial agreement).

THIS IS WHAT I WAS SAYING--SO (B) IS EMPIRICAL.

The problem does not arise if (B) holds and uninterpretable features delete in an "all or none" fashion, not selectively. In (44ii-iv), the full complement of $_features$ of T deletes the uninterpretable feature G of <there>, barring further raising. When EXPL raises to SPEC-T<def>, the probe (a single feature) deletes under matching as before, but G does not, because deletion requires matching with a full complement of $_features$ of the probe.

THIS IS THE KEY, OF COURSE.

Therefore successive-cyclic raising through SPEC-T<def> is possible.

{For a different approach, on the assumption that <there> has structural Case, see Lasnik (1995a). For a different perspective on a wide range of related issues, see Moro (1997).}

Reinterpretation of Attract in terms of Agree eliminates the need to introduce "checking domains." That is a step forward.

FINE, BUT REMEMBER WE'RE STILL USING THE NOTION MINIMAL DOMAIN TO DETERMINE EQUIDISTANCE. MINIMAL DOMAINS, OF COURSE, WHERE THE UNION OF CHECKING DOMAINS AND INTERNAL DOMAINS; INDEED, AT THE TIME WE DEFINED THEM, WE THOUGHT OF CHECKING DOMAINS AS 'THE REST', THEREFORE A SORT OF EMERGENT DOMAIN WHICH YOU HAVE GIVEN THE NATURAL NOTION 'MINIMAL DOMAIN' (UNDERSTOOD IN THE SENSE OF A POCKET OF L-RELATEDNESS REGULARITY WITHIN THE PROJECTION OF A CATEGORY) AND THE EQUALLY NATURAL NOTION 'INTERNAL DOMAIN' (UNDERSTOOD IN THE SENSE OF A POCKET OF THETA REGULARITY IN THE HEAD-COMPLEMENT CONFIGURATION).

The notion is complex, and furthermore unnatural in minimalist terms;

AS I SAY, THAT DEPENDS ON WHAT YOU MAKE OF MY PREVIOUS PARAGRAPH.

feature-checking should involve features, nothing more, and there is no simpler relation than identity. More important, the notion is irrelevant for the core cases: elements merge in checking domains for reasons independent of feature checking; and feature-checking takes place without dislocation to a checking domain.

THAT IS A WELL-TAKEN POINT, AND THE MAIN REASON WHY IT IS WORTH PURSUING THESE MATTERS.

As discussed, both properties are illustrated in expletive constructions. Much work on the topic has taken long-distance effects to be a property of these constructions, hence of an expletive-associate relation; various ideas have been explored as to how that relation is established. In chap. 4 of <MP>, a different approach is suggested: the long-distance effects are attributed to an T-associate relation that involves features only and is independent of the expletive. The reasons were theory-internal, but a broader range of cases adds empirical support. Long-distance effects are found without expletives in such constructions as (15i),

THOSE WERE INSTANCES OF QUIRKY DATIVE RAISING, SAY IN PSYCH CONSTRUCTIONS IN THE LANGUAGES THAT ALLOW THEM.

{Similar conclusions are supported by locative and quotative inversion (see Collins 1997), though with restrictions and complications, and similarities to other poorly understood constructions (e.g., "still unclear remain (are, seem to be) the answers to those questions").}

THESE ARE TRULY FASCINATING, SINCE OBVIOUSLY THEY DON'T INVOLVE AN EXPLETIVE, AND YET THEY PRESENT AGREEMENT-ASSOCIATE RELATIONS.

with EPP satisfied by raising of quirky Case; and expletive subjects are found without T-associate agreement when there is no accessible nominative.

THAT'S THE OTHER HALF OF THE COIN, SHOWING THAT EXPLETIVE/ASSOCIATE IS ONE THING, AND AGREEMENT/ASSOCIATE IS A DIFFERENT THING.

We return to some illustrations. The general conclusions are (45):

(45)(i) Long-distance agreement is a T-associate (probe-goal) relation

(ii) EPP can be satisfied by:

- (a) Merge of expletive
- (b) Merge of associate
- (c) Merge of _ closer to T than the associate

Case (a) is illustrated by T-associate agreement, with the definiteness effect.

YOU CAN IN FACT USE THAT EFFECT TO TEST WHETHER YOU HAVE THIS KIND OF RELATION, PARTICULARLY IF THE LANGUAGE DOESN'T HAVE OVERT EXPLETIVES. WE HAVEN'T SAID MUCH OF WHAT PREDICTS THE DEFINITENESS EFFECT, BUT WE'LL RETURN TO THAT.

Case (b) exhibits agreement of SPEC-T and T, but that is ancillary to the T-associate relation.

THAT IS, THE ASSOCIATE OF T HAS MOVED THERE FORE EPP REASONS.

In case (c) there is no definiteness effect and long distance T-associate agreement holds with embedded accessible nominative; or, if such an associate is lacking, T is default.

WE HAVEN'T SEEN THIS SORT OF EXAMPLE, I DON'T THINK, BUT IT WOULD BE ONE WHERE T MANAGES TO AGREE DOWN WITH SOME NOMINATIVE ELEMENT, AND A CLOSER ELEMENT HIGHER UP IN THE TREE, OR PERHAPS MORE STUFF IN THE NUMERATION, MERGES AT EPP LEVEL; THEN YOU WOULDN'T HAVE A DEFINITENESS EFFECT SINCE THERE IS NO EXPLETIVE. I'D SAY THAT LOCATIVE INVERSION IN ENGLISH, OR THOSE INSTANCE WE JUST SAW INVOLVING A PREPOSITIONAL PHRASE, MIGHT BE OF THE RELEVANT SORT (SO all of a sudden came Johnny with his gun). YOU MIGHT ALSO GET DEFAULT AGREEMENT, ALTHOUGH I'M NOT SURE HOW YOU GET RID OF THAT IF IT IS UNINTERPRETABLE.

More generally, we should not expect SPEC-head relations to have any special status. Within bare phrase structure, we cannot, for example, take the result of first Merge to _ to be sometimes a specifier and sometimes a complement, as in an X-bar-theoretic analysis that takes the object of _ to be its complement ("see John," "proud [of John]") but the subject of objectless _ to be its specifier (base forms of "John eat," "John proud").

WELL, OF COURSE, BUT THAT WAS NEVER REALLY TAKEN TO BE AN ISSUE BEFORE, ASSUMING THE HALE/KAYSER LINE FOR UNERGATIVES. I'M NOT SURE, EITHER, THAT THE ISSUE WAS EVER, FOR CHECKING DOMAINS, HEAD/SPECS, BUT RATHER 'WHAT REMAINS FROM MINIMAL DOMAINS ONCE YOU TAKE AWAY INTERNAL DOMAINS', SO A TOTALLY DERIVATIVE NOTION. THAT WAY WHAT YOU COULDN'T DISTINGUISH WAS HEAD ADJUNCTION FROM SPECS (UNIFORMITY AND SIMILAR CONSIDERATIONS DICTATING WHETHER SOMETHING HAD TO BE A HEAD OR NOT), BUT THAT WAS TAKEN TO BE A VIRTUE OF THE PREVIOUS SYSTEM.

The restriction to single specifier is also questionable: rather, we would expect first Merge, second Merge, etc., with no stipulated limit.

TRUE, BUT THAT TOO SEEMS ORTHOGONAL. WHETHER A SINGLE SPECIFIER EXISTS IS SOMETHING AKIN TO SAYING WHETHER THE SYSTEM SHOULD RECOGNIZE last Merge, AND IN ANY CASE IS AN EMPIRICAL ISSUE. WHICH IS ALL TO SAY, CONCEPTUALLY WE'VE SIMPLIFIED SOME THINGS, BUT IT'S NOT OBVIOUS THAT THE SYSTEM IS REMARKABLY BETTER. EMPIRICALLY, HOWEVER, WE HAVE IMPROVED IT IN QUITE INTERESTING WAYS, I BELIEVE.

I WON'T REPEAT THE CLARIFICATIONS I STARTED WITH IN THE PREVIOUS SECTIONS, WHICH BY NOW SHOULD BE ASSUMED. I WANT TO SAY, HOWEVER, THAT WHAT FOLLOWS IS BOTH SPECULATIVE AND HARD TO UNDERSTAND, WHICH MEANS I MAY BE TOTALLY WRONG IN MY INTERPRETATION--LET'S HOPE WE CAN COME UP WITH A SERIOUS ONE AS A GROUP.

6. Syntactic Objects

We are taking the elements that enter into derivations to be features and objects constructed from them in a restricted way:

- (46) (A) lexical items LI
(B) modified lexical items MLI
(C) sets K constructed from given elements a,b.

An MLI is an LI with uninterpretable features deleted.

OF COURSE, THE QUESTION HERE IS WHY THE SYSTEM BUILDS LI'S IN SUCH A WAY THAT THEY END UP BEING MLI'S ONLY AFTER CHECKING. IN OTHER WORDS, WHY ISN'T THE SYSTEM SMART ENOUGH TO MAKE THE LI'S WITHOUT THE UNINTERPRETABLE FEATURES. THERE WERE SOME SPECULATIONS WE SAW ABOUT THIS, BUT NONE IS EVEN A REAL CONJECTURE.

In case (C), K corresponds to the subtree dominated by a node in a standard phrase structure diagram; there are no objects corresponding to the nodes,

ANOTHER WAY OF SAYING THIS IS THAT THE SYSTEM WE'RE DEVELOPING DOESN'T TREAT TREES AS GRAPHS THAT YOU CAN MODIFY, OR (DIS)CONNECT HERE OR THERE, AND SO ON. THE REAL STUFF IS, ULTIMATELY, LI'S AND THEIR TRIVIAL COMBINATIONS. THE FORMAL OBJECTS WE REPRESENT AS $\{A, \{A,B\}\}$ ARE NOT LABELS FOR NODES, BUT THEY ARE, THEMSELVES, THE OBJECTS THE SYSTEM MANIPULATES. IN FACT, THE LABEL IN THE OBJECT I JUST GAVE YOU IS 'A', GIVEN THE FORMAL NOTATION, WHICH IS LIKE SAYING YOU'RE GIVING THE OBJECT FORMED BY ASSOCIATING A AND B A'S COLOR (FOR THESE ISSUES, SEE MY R&R, ESPECIALLY THE APPENDIX BY NUNES AND THOMPSON, WHERE THESE NOTIONS ARE EXPLICITLY DEFINED).

and no non-branching projections.

THAT IS, UNDER THE ASSUMPTION THAT OBJECTS ARE OBTAINED ONLY UNDER MERGE. THIS IS A NON-TRIVIAL ASSUMPTION, ACTUALLY, BOTTOMING OUT IN A KIND OF PROBLEM WHICH MIGHT BE MORE SERIOUS THAN IT SEEMS. FORMALLY, TERMS ARE DEFINED AS (I) SYNTACTIC OBJECTS (E.G. K ABOVE), THIS IS THE BASE STEP OF THE DEFINITION; AND (II) MEMBERS OF MEMBERS OF TERMS (E.G. A AND B IN THE LITTLE CASE I GAVE YOU). IN TURN, YOU CAN DEFINE LABELS AS TERMINALS WHICH ARE NOT TERMS. FINE, BUT THEN GO TO

THE BOTTOM OF THE OBJECTS: WHAT IS, SAY, B? IS IT A TERM? YOU CLEARLY WANT THAT, SINCE B IS MANIPULATED BY MERGE, MOVE, AND SO FORTH. BUT THEN IT CAN'T BE A LABEL; ALRIGHT, SO WHAT IS ITS LABEL? INTUITIVELY, YOU WANT IT TO BE 'B', BUT WE'VE JUST ESTABLISHED THAT 'B' IS A TERM, HENCE CAN'T BE A LABEL. THIS SORT OF PROBLEM ARISES WITH ALL TERMINALS, ACTUALLY. IT'S AS IF YOU WANTED TO SAY THAT THE TERMINAL IS REALLY '{B}', WHILE THE LABEL IS 'B'. BUT THAT WOULD BE, IN EFFECT, A NON-BRANCHING PROJECTION, AND BESIDES: HOW DO YOU JUSTIFY GETTING THAT SORT OF OBJECT IN THE SYSTEM? YOU MAY THINK THAT THIS IS JUST TERMINOLOGY, BUT IT'S NOT THAT SIMPLE, BECAUSE MINIMALISM SUGGESTS GETTING RID OF CODING TRICKS, AND THE ISSUE IS THEN: IF YOU *REALLY* GET RID OF CODING TRICKS, WHAT IS THE ULTIMATE CODE? HOW DO I REALLY KNOW AN 'A' FROM A 'B'; OR WORSE STILL: *WHAT* IS AN 'A' OR A 'B'? THINK OF THIS FROM THE POINT OF VIEW OF A NATURAL SYSTEM. WHAT DOES IT MEAN TO SAY THAT SUCH AND SUCH A CELL OR MOLECULE OR WHATEVER IS OF TYPE 'A' OR TYPE 'B'. TYPICALLY, WE DO THAT IN TERMS OF A CLASSIFICATION OF A CERTAIN BEHAVIOR, VIS-A-VIS A GIVEN THEORETICAL DESCRIPTION. BUT ULTIMATELY, REALISTICALLY AS IT WERE, TYPE 'A' OR 'B' IS ROOTED IN SOME KIND OF BIO-CHEMICAL, OR PHYSICAL PROPERTY, OR SOME SUCH THING. WE DON'T THINK, THESE DAYS ANYWAY, OF HEAT, SAY, AS A MERE QUALITY OF BODIES; WE THINK WE UNDERSTAND HEAT WHEN WE RELATE IT TO A KIND OF ENERGY, SAY MOLECULE MOTION. THINGS GET WAY MORE ARCAINE AS LEVELS OF REALITY GET SIMPLER OR MORE COMPLEX, AND THUS THE PROPERTY OF COLOR, CONSIDERED BY A QUANTUM PHYSICIST, IS A BIZARRE CHARACTERISTIC OF QUARKS, A SO-CALLED 'QUANTUM NUMBER'. IS THAT NUMBER A 'MERE' NOTATION OR SOMETHING WHICH, AT SOME LEVEL, WILL BE AS TRIVIAL AS COLOR WAS FOR STANDARD BODIES? PERHAPS, THOUGH NOT WITHIN THE CONFINES OF PRESENT-DAY STORIES. THE POINT IS TWOFOLD. FIRST, VERY BASIC NOTATIONS ALWAYS BOTTOM OUT IN A BIZARRE, SEEMINGLY ARTIFICIAL WAY. BUT SECOND, THIS IS NOT TO SAY THAT WE SHOULD SHY AWAY FROM THIS KIND OF PROBLEM, SINCE WE MAY ACTUALLY MAKE PROGRESS BY RESOLVING PUZZLES AND CONTRADICTIONS OF THE SORT JUST MENTIONED. I WILL RETURN LATER ON TO THE ISSUE OF THE LABELS OF TERMINALS, AND EVEN MORE BASICALLY, TO THE NATURE OF LABELS THEMSELVES (SUCH BASIC QUESTIONS AS WHY WE HAVE 'V' AND 'N' DIMENSIONS, AND SO ON).

Theta structure and similar semantic roles are based on pure Merge of XP to substantive LIs or their projections. Checking Theory, in contrast, involves uninterpretable features of functional categories and is reduced to feature matching under conditions (40). The two theories differ in nature and implementation -- plausibly, since they are conceptually quite distinct.

I ACTUALLY AGREE WITH THIS WHOLEHEARTEDLY, WHICH IS WHY I FIND IT A MISTAKE TO CONFLATE THE TWO SORTS OF INFORMATION IN A SINGLE SYSTEM. INDEED, I EVEN THINK THAT A SINGLE COMPONENT TO CAPTURE BOTH OF THESE IS PLAUSIBLY TOO COARSE GRAINED.

Something like Theta Theory is a property of any language-like system, while Checking Theory is specific to human language, motivated (we are speculating) by legibility conditions.

TAKING THIS REMARK SERIOUSLY, I WONDER WHETHER CHOMSKY WOULD GO AS FAR AS TO ADMITTING THAT THETA THEORY MAY BE A PROPERTY OF A MIND THAT DOESN'T HAVE CHECKING THEORY. I WOULDN'T FIND IT IMPLAUSIBLE THAT IN A STAGE OF OUR EVOLUTION WE HAD NO CHECKING THEORY. THIS IS ARGUABLY IMPORTANT, SINCE SOME CURRENT SPECULATIONS BY BICKERTON AND HIS ASSOCIATES (CHECK THEIR WEB PAGE IN THE MIT PRESS SITE) WANT TO BLAME THE EVOLUTION OF LANGUAGE ON A 'BALDWIN EFFECT' ON THETA THEORY: YOU GOT THETA THEORY, AND THEN THE REST SORT OF FELL IN PLACE... I FIND THAT VERY HARD TO BELIEVE, AND WOULD RATHER EXPECT A CRUCIAL MUTATION TO INTERVENE BETWEEN THOSE TWO STAGES (WHICH IS NOT TO SAY, OF COURSE, THAT THE MUTATION GAVE US CHECKING THEORY *DIRECTLY*; IT COULD HAVE GIVEN US SOMETHING FUNCTIONALLY UNRELATED, OF COURSE).

The Chain Condition is an expression of this duality, and should fall out as a descriptive observation, along with other properties of chains.

Footnote{Among these, ECP, subjacency, and other conditions. But many problems remain. See references of note UR1A.}

THIS REFERS TO THE FOOTNOTE ABOUT ISLANDS. NEEDLESS TO SAY, THE TRICK HERE IS TO RELATE WELL-KNOWN PROPERTIES OF CHAINS (SUCH AS THE OBSERVATION THAT THE TAIL OF THE CHAIN GETS THE THETA ROLE AND THE HEAD DOES THE CHECKING) TO DERIVATIONAL DYNAMICS OR NATURAL INTERFACE CONDITIONS, AS USUAL.

Apart from Merge of selected XP,

AND REMEMBER THAT FOR CHOMSKY THIS INCLUDES EPP DEMANDS, WHICH I BELIEVE IS NOT A VERY USEFUL NOTATION.

narrow syntax involves only feature deletion to form reduced MLIs, sometimes associated with Merge in non-theta positions yielding dislocation.

According to this conception, agreement (hence movement) is driven by uninterpretable features of the probe, which must be deleted for legibility. The operation Greed of MP, in contrast, was driven by uninterpretable features of the goal. The principle we are now entertaining is what Lasnik (1995a,b) calls Enlightened Self-Interest,

ACTUALLY, AS I UNDERSTOOD SELF-INTEREST, IT WOULD ALLOW AN OPERATION IN TERMS OF A PROBE OR GOAL TRIGGER, NOT JUST A PROBE TRIGGER.

with the further requirement that the matched probe delete: we might call the principle Suicidal Greed. Suicidal Greed does not have the "look-ahead" property of Greed, a complexity reduction that could be significant, as discussed.

With this shift of perspective, structural Case is demoted in significance. The Case Filter still functions indirectly in the manner of Vergnaud's original proposal, to determine the distribution of noun phrases. But what matters primarily are the probes, including f-features of T, v. That reverses much of the recent history of inquiry into these topics and also brings out more clearly the question of why Case exists at all. The question arises still more sharply if matching is just identity, so that Case can never be attracted; operations are not induced by Case-checking requirements. Recall that lexical category also cannot be attracted and does not induce operations, raising the same questions.

ALTHOUGH HERE IT IS LESS CLEAR THAT ANY OBVIOUS EMPIRICAL CONCLUSION FOLLOWS. AND NOTE:

Footnote {Perhaps substantive lexical categories do not exist, only bare roots.

WHICH WOULD BE CONSISTENT WITH THEIR NOT BEING INVOLVED IN CHECKING NOT HAVING MUCH OF AN EMPIRICAL CONSEQUENCE (UNLIKE WITH CASE). CONCEPTUALLY, ALSO, IT WOULD PARTLY ANSWER OUR CONCERNS ABOUT LABELS ABOVE, THOUGH MANY QUESTIONS STILL REMAIN.

Configurational and morphological properties, along with interpretable noncategorical features of the root, would then determine relevant structural properties, as in Semitic.

THE SEMITIC CASE IS COMPLEX, BECAUSE IN SOME ANALYSES (MACARTHY'S) IT IS NOT THAT THE ROOT, PER SE, HAS THE KEY PROPERTIES, BUT RATHER THAT THE VOWELS AND CONSONANTS ARE IN DIFFERENT TIERS, SO THAT NOTIONS LIKE ROOT AND SO ON MAKE RELATIVELY LITTLE SENSE. THE FACT IS THAT DEPENDING ON VOWEL COMBINATION IN SEMITIC YOU GET DIFFERENT TENSES, ASPECTS, AND SO ON. IN ANY CASE, THE GENERAL POINT THAT THE LEXICAL CATEGORY IS JUST A SYNTACTIC FORMATIVE OF SOME SORT IS ATTRACTIVE, AND RATHER WELL MOTIVATED IN A VARIETY OF LANGUAGES:

The possibility is suggested by work since the '60s on derivational vs. transformationally-induced morphology. See Marantz (1997) for pertinent discussion.}MAR

I PERSONALLY SUSPECT, HOWEVER, THAT THERE MAY BE A PARAMETER HERE, ALONG THE LINES HINTED AT BY LASNIK IN HIS RECENT WORK (SEE THE COLLECTED ESSAYS, MINIMALIST ANALYSIS (1998)). MY REASONS FOR THIS ARE EMPIRICAL, AND BUILD ON LASNIK'S FINDINGS. THUS, PORTUGUESE DIFFERS FROM ENGLISH IN ALLOWING ELLIPSIS ONLY WHEN TOTAL IDENTITY OF MATCHING HOLDS. IN PORTUGUESE, FOR INSTANCE, YOU CANNOT SAY John will visit Mary and Peter has already, ALTHOUGH YOU CAN SAY John will visit Mary and Peter

will also. IN A SENSE, PORTUGUESE IS THE EXPECTED INSTANCE, IF ELLIPSIS INVOLVES SOME SORT OF IDENTITY REQUIREMENT, FOR AFTER ALL IN ENGLISH THE PRONOUNCED VERB *visit* IS NOT THE SAME AS THE ELLIDED ONE *visited*, IN THE FIRST INSTANCE. HOWEVER, IT COULD BE THAT, AS LASNIK SUGGESTS, THE ENGLISH VERB IS JUST THE *visit* PART, NOT THE *-ed* PART, WHICH MIGHT APPEAR IN THE COURSE OF THE DERIVATION (IN YOUR FAVORITE WAY). THAT WOULD PRESERVE IDENTITY. BUT THEN THE PORTUGUESE VERB MUST HAVE THE EQUIVALENT OF *visit* VS. *visited* *ALREADY IN THE NUMERATION*. IS THE DIFFERENCE ONE OF CATEGORIAL FEATURES, EXPRESSED BY THE RELEVANT MORPHEME? PERHAPS. IF SO, THOUGH, IS IT THE CASE THAT THE ENGLISH VERB DOES NOT HAVE THE RELEVANT CATEGORIAL MORPHEME UNTIL IT PICKS IT UP? PERHAPS. DEEP ISSUES THAT SHOULD BE SERIOUSLY PURSUED.

For Case, a plausible answer is the one already suggested: uninterpretable features activate the goal of a probe, allowing it to implement some operation (Agree or Move).

Footnote {A different motivation, based on the need to overcome ambiguity in the composite verbal element at LF, is developed by Uriagereka (1996).

THIS IS NOT REALLY ACCURATE. THE MOTIVATION FOR HAVING CASE AT ALL IS NOT WHAT CONCERNED ME THERE, AND IN FACT I AM PERFECTLY WILLING TO ADMIT CHOMSKY'S PRESENT MOTIVATION, OR EVEN A THIRD POSSIBILITY, THAT THE MOTIVATION IS STILL A DIFFERENT ONE (WHATEVER IT IS), AND CHOMSKY'S CONDITION FOLLOWS AS A SIDE EFFECT. BUT WHAT I WAS INTERESTED IN WAS THE ISSUE OF WHY THERE ARE *DIFFERENT* CASES, WHY NOT A SINGLE MARK FOR ALL INSTANCES. IN FACT, WHY THE CASES DIFFER WITHIN THE DOMAIN OF WHAT WE NOW SEE AS THE CYCLIC ACCESS TO A NUMERATION: WHY IS THE GRAMMAR CONCERNED WITH MARKING THINGS FOR DIFFERENCE WITHIN THOSE DOMAINS. THEN THE POINT IS NOT REALLY AMBIGUITY, BUT SOMETHING MORE SERIOUS. SAY THE NUMERATION IS A SET, AND YOU HAVE, WITHIN A GIVEN CHUNK THAT YOU ACCESS, TWO [D] ELEMENTS. HOW DO YOU MARK THAT? IF THEY ARE DIFFERENT, SAY [a] VS. [the], AND IF YOUR GRAMMAR SEES, ALREADY AT THIS LEVEL, THE PHONETIC AND SEMANTIC MATRIX, THEN YOU'RE FINE. ON THE OTHER HAND, IF EITHER OF THOSE TWO PREMISES FAILS, AND FOR INSTANCE YOU HAVE TWO [the]S, THERE'S NO WAY OF REPRESENTING THIS OTHER THAN WITH A CODING TRICK. OF COURSE, IT IS A TRICK IF YOU DO IT: MARKING THEM 1 VS. 2, OR SOMETHING LIKE THAT. BUT IF THE GRAMMAR DOES IT, IT MAY BE A PROPERTY; PERHAPS THE GRAMMAR JUST DOES THAT: USING +NOMINATIVE VS. -NOMINATIVE IN ORDER TO ENSURE THAT THE DIFFERENT TOKENS EXIST IN THE NUMERATION. ACTUALLY, I MADE THE REASONING FOR LF BECAUSE AT THE TIME I WROTE THE ARTICLE THAT WAS THE POINT WHERE DOMAINS WERE FORMED (WE DIDN'T HAVE CYCLIC ACCESS TO THE NUMERATION). IN PRESENT TERMS YOU CAN MAKE THE POINT MUCH MORE DIRECTLY AND NATURALLY BY WAY OF THE NUMERATION. FRANKLY, EVEN IF ONE GOES WITH CHOMSKY'S GENERAL REASONING THE PRESENT REASONING STILL HOLDS, AND HE'S ADMITTED THAT

MUCH IN A COUPLE OF TALKS.

One can also think of various functional arguments: the familiar trade -off between order rigidity and richness of inflection; facilitation of search for attracted elements, etc.}

THE FIRST OF THESE I DON'T UNDERSTAND AS A *MOTIVATION*, ALTHOUGH I ADMIT IT IS A PLAUSIBLE SIDE-EFFECT (EVEN THOUGH YOU HAVE PROBLEMATIC INSTANCES). THE SECOND IS AGAIN A NICE SIDE-EFFECT, BUT IT IS HARD TO SEE HOW CASE FEATURES WOULD EMERGE IN THE SYSTEM FOR THAT REASON.

It follows that after deletion of structural Case of DP, the phrase cannot move further to an A-position and its f-set cannot induce deletion (though it is still "visible" to a probe, allowing defective intervention effects as in (42)). Suppose quirky Case is (theta-related) inherent Case with an additional structural Case feature, as often suggested in one or another form. Then it too is immobile once it reaches a Case-checking position.

THIS, OF COURSE, IS A CONJECTURE ABOUT QUIRKY CASE. IN ESSENCE, EVERY TIME CHOMSKY SEES A FREEZING EFFECT, HE'LL INVOKE ONE OF THESE UNINTERPRETABLE FEATURES THAT DOES NOT CORRESPOND TO AN IDENTICAL FEATURE IN THE PROBE.

Footnote {Pure inherent Case we take to be a distinct phenomenon, "invisible" to matching, as if inherent Case inactivates the f-set.

THAT IS A CURIOUS SPECULATION, SUGGESTING THAT THE f-SET IS A FORMATIVE AS WELL (RECALL THE DISCUSSION ABOUT CATEGORIAL FEATURES), FOR OTHERWISE HOW CAN YOU TARGET IT FOR "INVISIBILITY"? NEEDLESS TO SAY, IF MANY OF THE FEATURES ARE REAL IN THE OBJECTUAL SENSE, SO MUCH SO THAT YOU CAN TARGET THEM FOR CERTAIN OPERATIONS, THE QUESTIONS OF A) WHY THEY BUNDLE UP INTO WORDS AND B) WHY THEY HAVE A SPECIAL STATUS HIT US RIGHT BACK. THE MATTER ALSO TOUCHES ON FUNDAMENTAL WORRIES ABOUT ATOMISM AND DECOMPOSITION THAT I PLAN TO RETURN TO.

See note LIGHT, and for more on the topic, McGinnis (1998) and sources cited.}

THIS IS A REFERENCE TO FOOTNOTE 31, AND IS NOT PARTICULARLY USEFUL.

If the f-features of T that check the structural Case of raised quirky Case themselves delete, we have default T;

I'M NOT SURE WHAT CHOMSKY MEANS BY DEFAULT T. THIS MIGHT BE A TYPO, AND HE MIGHT HAVE MEANT DEFAULT AGREEMENT. PARTICULARLY SINCE:

if they remain, we have remote agreement with some lower accessible nominative.

THE WAY I UNDERSTAND THIS, REMOTE AGREEMENT SHOWS UP WITH A LOWER PHRASE, WHICH IS THE ONE WHICH WILL DELETE THE T FEATURES WHICH HAVE NOT BEEN DELETED IN RELATION TO THE QUIRKY PHRASE. IF, ON THE OTHER HAND, THE FEATURES OF T DO DELETE, THEN THE KIND OF AGREEMENT THAT SHOWS UP IS NOT WITH A DISTANT PHRASE, BUT IS MERELY DEFAULT (I TAKE IT THAT THAT'S WHAT 'DEFAULT T' MEANS). OF COURSE, IN SOME INSTANCES YOU MIGHT NOT BE ABLE TO TELL DEFAULT AGREEMENT FROM, SAY, THIRD PERSON SINGULAR AGREEMENT--BUT THAT'S AN ACCIDENT. YOU MIGHT BE ALSO WONDERING WHY, WHEN THE T FEATURES *ARE* DELETED BY THE RAISED QUIRKY ELEMENT NO AGREEMENT HOLDS, AND YOU GO INTO DEFAULT. HERE'S WHERE THE IDEA THAT QUIRKY CASE MAKES THE f-SET INVISIBLE HAS SOME EMPIRICAL BITE. WHICH AS I SAID MUST MEAN THAT, SOMEHOW, YOU CAN ACCESS THAT SET AND 'TURN IT OFF'.

The descriptive observations seem generally accurate. In these terms, the Visibility Thesis and Chain Condition reduce to LF convergence.

In his detailed review of Icelandic agreement, Sigurdsson (1996) concludes that remote NOM allows number agreement but not first/second person agreement.

THIS, INCIDENTALLY, IS SLIGHTLY WORRISOME, FOR AS WE SAW WE MUST DISTINGUISH REMOTE AGREEMENT FROM DEFAULT AGREEMENT. NOW IT SEEMS AS IF IT REDUCES TO A MATTER OF SINGULAR VS. PLURAL. FOR WHAT IT'S WORTH, I SUSPECT THAT SIGURDSSON'S DISCOVERY DOES NOT EXTEND TO OTHER LANGUAGES, FOR INSTANCE THE ROMANCE ONES--WHICH IS GOOD FOR THE PURPOSES OF THE ANALYSIS CHOMSKY IS SUGGESTING.

That would follow if the [person] feature of T reduces to [3person] (the default choice) when it attracts quirky Case or EXPL to SPEC-T.

AGAIN, NOT OBVIOUSLY FOR ROMANCE. BESIDES, IT'S NOT OBVIOUS WHY THIS SHOULD BE. I SUPPOSE CHOMSKY IS TRYING TO RELATE THE IDEA TO HIS CONJECTURES ABOUT EXPLETIVES HAVING A PERSON FEATURE, BUT THAT IS SOMEWHAT DUBIOUS AS WELL. IN ANY CASE, THIS IS ALL MACHINERY.

Footnote {Reformulation is needed if [3person] is lacking, as has been suggested.

THIS COULD BE MEANT RADICALLY: IF [3person] IS NOT A GRAMMATICAL NOTION. THERE'S FAIRLY GOOD EVIDENCE FOR THIS, GOING BACK, IN FACT, TO OLD STUDIES BY BENVENISTE. OF COURSE, IF THAT WERE TRUE, AND THE SYSTEM IS, RATHER, SOMETHING LIKE +/-SPEAKER (I.E. FIRST VS. SECOND PERSON) VS. *NO* SPEAKER REPRESENTATION, THEN THE FEATURE IN QUESTION MUST BE A DIFFERENT ONE ALTOGETHER. HOWEVER, CHOMSKY DOESN'T SEEM TO HAVE IN MIND ANYTHING THIS RADICAL. RATHER, HE SEEMS TO BE SPEAKING OF SOME KIND OF 'PERSON REDUCTION', WHICH I DON'T REALLY

UNDERSTAND, WHEN (AS I FOLLOW THIS) THE RELEVANT GRAMMATICAL RELATION HOLDS ACROSS TWO DIFFERENT CLAUSES:

Sigurðsson concludes that third person NOM requires agreement in the monoclausal construction [DAT V-T [vp t NOM]] and allows it as an option, alternating with default T, in the biclausal construction [DAT V-T t [TP NOM...]] (t the trace of DAT; irrelevant properties omitted).

I GUESS! I HAVE A HARD TIME INTERPRETING EVEN THE *RELEVANT* PROPERTIES. I ASSUME THE ISSUE IS THAT, FOR SOME REASON, THE PERSON FEATURE IS COMPLETELY KNOCKED OUT, OPTIONALLY, IN THE SECOND INSTANCE INVOLVING EMBEDDING, FOR REASONS THAT CHOMSKY DOESN'T UNDERSTAND:

These facts could be expressed in terms of optionality of [person]-reduction in matrix T, leaving open the reasons.}PERS

Suppose expletive EXPL is merged in SPEC-T without movement. Agreement is manifested for it-type EXPL with a full complement of f-features, and there-type EXPL becomes inactive, indicating that its uninterpretable feature deletes ([person], we have assumed).

THIS IS INTENDED AS THE KEY DIFFERENCE BETWEEN it AND there. IN ESSENCE, CHOMSKY IS SAYING THAT it HAS ITS f-SET INTACT, UNLIKE there. BUT HOW CAN THIS BE DETERMINED? TRUE, there SHOWS 'AGREEMENT WITH THE ASSOCIATE' (SEE BELOW ON WHAT THIS MEANS). HOWEVER, COULDN'T WE SAY SOMETHING SIMILAR FOR it? ISN'T THE ASSOCIATE OF it A WHOLE CLAUSE? THE EVIDENCE GOES BOTH WAYS. ON THE ONE HAND YOU HAVE EXAMPLES INVOLVING 'LONG DISTANCE ACCESSIBLE SUBJECTS' LIKE they think that it is likely that pictures of each other will be on sale (SEE THE CHAPTER ON BINDING IN LASNIK AND URIAGEREKA). THESE COULD BE TRIVIALY ANALYZED IF, AT LF, THE REPRESENTATION IS ALONG THE LINES OF they think that [that pictures of each other will be on sale] is likely. (ALTHOUGH WE'VE ALREADY SEEN SERIOUS TROUBLES WITH HAVING BINDING BLINDLY APPLY AT LF.) ON THE OTHER HAND, McCLOSKEY NOTED PROBLEMATIC EXAMPLES ALONG THE FOLLOWING LINES: [That he will lie again and that he won't be prosecuted] are likely to happen, VS. It is/*are likely to happen that he will lie again and that he won't be prosecuted. BASICALLY, IT SEEMS THAT AGREEMENT IS DETERMINED BY it, NOT THE ASSOCIATE.

But these are properties of Agree, not Merge.

Footnote {If Merge could induce agreement, vP-internal subject would manifest object agreement and have accusative Case, which would delete, preventing raising to SPEC-T.}

I'M NOT SURE I UNDERSTAND THIS FOOTNOTE. TRUE, IN THE DERIVATION CHOMSKY OUTLINES, WE GET A CRASH (EITHER THE FEATURE OF T IS NOT CHECKED OR WE VIOLATE THE MLC). BUT SUPPOSE MERGE CAN INDUCE AGREEMENT; WHAT FORCES IT? THE GRAMMAR WOULD FACE AN OPTION

WITHIN VP, EITHER AGREEMENT-VIA-MERGE WITH SUBJECT OR AGREEMENT-VIA-MOVEMENT WITH OBJECT (THE STANDARD CASE). OF COURSE, THE FIRST INVOLVES LESS OPERATIONS, BUT THIS IS NOT A METRIC OF ECONOMY, AT LEAST NOT IN THE TERMS OF CHAPTER 4, SINCE THE RESULT IS NOT CONVERGENT. SO EITHER CHOMSKY IS ENTIRELY GIVING UP THAT SYSTEM (AND IT IS NOT CLEAR TO ME WHY, SINCE THE LOOK-AHEAD THAT THE SYSTEM INVOLVES IS REALLY TRIVIAL ONCE WE HAVE CYCLIC ACCESS TO NUMERATIONS), OR ELSE THE REASONING IN THE FOOTNOTE DOESN'T HOLD, UNLESS I'M MISSING IT. OF COURSE THAT DOESN'T MEAN THE CONCLUSION IS WRONG. FOR WHAT IT'S WORTH, THOUGH, NUNES AND RAPOSO ARGUE ON EMPIRICAL GROUNDS THAT AGREEMENT CAN HAPPEN UNDER MERGE IN SOME INSTANCES INVOLVING PORTUGUESE PARTICIPIALS. AT ANY RATE, JUDGING FROM (47iii) BELOW, I THINK CHOMSKY MEANT WHAT HE SAID.

The results are expected if EXPL is an X₀ head and its [person] feature is uninterpretable, therefore able to probe its domain T' (= D(EXPL)), locating the f-set of T as the closest goal.

MARK HIS WORDS. HE'S SAYING THAT SINCE it IS A PRONOMINAL ELEMENT, IT IS BOTH A HEAD AND A MAXIMAL PROJECTION. AS MAXIMAL PROJECTION IT CAN BE IN A SPEC, SATISFYING SOME EPP REQUIREMENT. BUT AS A HEAD, IT IS ABLE TO PROBE ITS DOMAIN, WHICH TURNS OUT TO BE T' (TECHNICALLY THIS IS PROBLEMATIC IF LABELS ARE PRIMITIVE AND IF WE TAKE DOMAINS TO BE COMPLEMENT DOMAINS, FOR AFTER ALL T' IS NOT THE COMPLEMENT OF THE EXPLETIVE; HOWEVER, HOLD YOUR JUDGEMENT ON THIS, SINCE I WILL SHOW YOU THE SORTS OF RESULTS THAT H&U CAN ACHIEVE BY SUSPENDING DISBELIEF WITH REGARDS TO THIS MATTER). OF COURSE, THE f-SET OF T IS FOUND IMMEDIATELY.

The uninterpretable probe deletes, and the f-set of T as well if EXPL has a full f-set.

THERE IS MUTUAL DELETION HERE BECAUSE BOTH SETS OF FEATURES ARE UNINTERPRETABLE, AND BY ASSUMPTION it HAS A FULL SET OF FEATURES, UNLIKE there (THIS IS IMPORTANT, OR THE SYSTEM WOULD SKIP THAT STEP, AS SEEN BEFORE FOR there).

We therefore have agreement but not via merger: full or partial agreement depending on the f-set of EXPL, which becomes inactive.

Footnote{Questions remain about French-style *il*, with a definiteness effect and other properties.

THE DEFINITENESS EFFECT IS IMPORTANT BECAUSE IT SUGGESTS, AGAIN, AN ASSOCIATE. OF COURSE, EVERY TIME YOU HAVE AN ASSOCIATE YOU DON'T REALLY HAVE SOMETHING ALONG THE LINES OF THE DIRECT AGREEMENT WE HAVE JUST CONSIDERED.

Suppose the [person] feature of T can delete or shift to default with there-type EXPL, in the

manner suggested for quirky Case.

WELL, THERE WASN'T MUCH OF A *MANNER* SUGGESTED FOR QUIRKY CASE, BUT A SUGGESTION THAT PERHAPS THIS SORT OF CASE ENTAILS THE DELETION OF THE [person] FEATURE. SINCE I DON'T KNOW HOW THIS HAPPENS (IS IT THROUGH MERGE, FOR INSTANCE?), THEN I DON'T KNOW HOW TO TAKE THE SUGGESTION. SAY FOR CONCRETENESS THAT IF A MERGES TO B, AND B HAS A PERSON FEATURE, AND A HAS WHATEVER EXPLETIVES AND QUIRKY CASE SHARE IN COMMON (?)--CALL THAT THE ? FEATURE--, THEN B'S PERSON FEATURE DELETES, IN THE SENSE OF BEING MASQUED FOR INTERPRETATION. THEN WHAT? YES:

That might accommodate "list readings" (as in "there (is/*am, remains/*remain) only me," "there (are, remain) only us (John and Bill)," in response, say, to "who's still here to do the work?"), with no person agreement or nominative Case assignment, in the absence of the full complement of f-features.

BUT HOW DOES THIS RELATE TO THE FRENCH *il*? WAS THIS MEANT AS A SUGGESTION FOR THAT ELEMENT? I READ THE SYNTAX IN THE FOOTNOTE THAT WAY, BUT HAVE A HARD TIME SEEING HOW THE IDEA, THOUGH CLEARLY INVOLVED, RESOLVES THE FACT THAT FRENCH *il* LOOKS LIKE ENGLISH *it*, YET BEHAVES LIKE ENGLISH *there* IN INDUCING A DEFINITENESS EFFECT OR LOCALITY EFFECTS WITH AN ASSOCIATE, YET LIKE ENGLISH *it* IN AGREEMENT TERMS OR IN TERMS, APPARENTLY, OF ANAPHORIC LICENSING (THOUGH THAT IS ALL RATHER DUBIOUS, AS WE ALREADY SAW).

Also to be accounted for is the fact that in v-phases the external argument pronoun does not agree with the v head, possibly indicating that argument pronouns have true D-N structure, unlike expletives.

THAT IS, he BEHAVES JUST LIKE the man WHEN INSIDE vP AS A SUBJECT, AND THIS SHOULDN'T BE THE CASE IF he WERE A HEAD. BUT PERHAPS IT ISN'T, CLEARLY SO IF POSTAL WAS RIGHT IN HIS INITIAL INSIGHT THAT he REALLY STANDS FOR the one.

See Cardinaletti and Starke (1994), Uriagereka (1988) for some relevant considerations. These remarks, needless to say, barely touch on a rich array of questions.}

The head of an A-chain can undergo A'-movement, of course, with different features accessed. Take wh-movement. This would be point-by-point analogous to A-movement if the wh-phrase has an uninterpretable feature [wh-] and an interpretable feature [Q], which matches the uninterpretable probe [Q] of a complementizer in the final stage; successive-cyclicity could then function in the manner discussed.

Footnote{To complete the analogy, C (and v with its f-set deleted) may have a nonspecific P-feature analogous to [person] for Tdef, perhaps contingent on assignment of the EPP-feature to

a phase; see discussion of (24), (25).}

THIS MIGHT ACTUALLY RELATE TO THE LITTLE PSEUDO-AGREEMENT ELEMENTS ONE SEES IN SOME LANGUAGES IN SUCCESSIVE CYCLIC WH-MOVEMENT. SOME OF THESE JUST SHOW UP INDIRECTLY, AS IN SPANISH WHERE TORREGO'S VERB PREPOSING SEEMS TO BE INVOLVED. BUT IN OTHER LANGUAGES WHAT SHOWS UP IS MORE RADICAL, FOR INSTANCE A WH-HEAD IN BASQUE. THESE, AS FAR AS I KNOW, DIVIDE IN TWO. IN BASQUE AND JAPANESE, SAY, THE HEAD IN QUESTION SHOWS UP IN THE LAST CYCLE, THE ONE WHERE WH-INTERPRETATION OBTAINS--SO IT COULDN'T BE A PHASE-DEPENDENT MATTER, AT LEAST NOT TRIVIAALLY. IN OTHER LANGUAGES, HOWEVER, THE HEAD APPEARS IN ALL CYCLES. CHAMORRO IN CHUNG'S ANALYSIS IS THE MOST OBVIOUS INSTANCE, ALTHOUGH PERHAPS PARTIAL WH-MOVEMENT IN SOME GERMANIC DIALECTS CAN ALSO BE SEEN IN THAT LIGHT (THOUGH NOT OBVIOUSLY, SINCE THE LATTER TENDS TO BE OPTIONAL AND, IF HERBURGER IS RIGHT IN HER 95 ANALYSIS, IT INVOKES A SLIGHTLY DIFFERENT READING). IN THE CHAMORRO CASE, AT LEAST, IT IS CERTAINLY SOUND TO SAY THAT AGREEMENT IS IN THE VARIOUS C PROBES, AND RELATES TO A PHRASE MADE VISIBLE IN TERMS OF A CASE-ANALOGOUS WH-ELEMENT.

The wh-phrase is active until [wh-] is checked and deleted.

THIS IS THE CASE-ANALOGUE, WHICH YOU MAY RECALL IS DIFFERENT FROM WHAT WE SAW IN CHAPTER 4, WHERE THE WH-FEATURE IS NOT DELETED.

The wh-island condition is then a defective intervention effect of the type (42):

THAT WAS AN INSTANCE OF THE SORT $\alpha \beta G$, WHERE α IS COMMAND.

the [Q] feature of the already checked wh-phrase (β in (42)) bars attraction of lower [Q], but cannot move or check the uninterpretable feature of the probe.

I REMAIN SKEPTICAL, FOR THREE REASONS: (I) THE INTERVENTION EFFECT IN THIS INSTANCE IS VERY MILD, COMPLETELY ATTENUATED IN SOME INSTANCES NOTED BY ROSS (INVOLVING MOVEMENT ACROSS TENSELESS CLAUSES); (II) THE INTERVENTION EFFECT HAS TO BE PARAMETERIZED (RIZZI EFFECTS), AND IT IS HARD TO SEE HOW TO PARAMETERIZE THE LOCALITY EFFECT, PARTICULARLY SINCE IN LANGUAGES WHERE YOU CAN MOVE OVER ANOTHER WH-PHRASE YOU NONETHELESS HAVE STANDARD LOCALITY LIMITATIONS ON A-RELATIONS; AND (III) THE INTERVENTION EFFECT ARISES EVEN IN INSTANCES WHERE WH-MOVEMENT IS NOT OBVIOUS (LASNIK AND SAITO SHOW VARIOUS EXAMPLES OF THIS SORT).

A possible analysis of wh- in situ constructions is that the [wh-] feature pied-pipes only the head (overtly or covertly).

Footnote{Following ideas of Watanabe (1992), Hagstrom (1998).

ALTHOUGH THE QUESTION STILL REMAINS OF WHY WHEN SOMETHING HAS ALREADY MOVED OVERTLY, AND ONLY THEN, NOTHING ELSE MOVES OVERTLY-OF COURSE IN RELEVANT LANGUAGES, SINCE IN OTHERS THIS IS PRECISELY WHAT HAPPENS.

This is not to be confused with the distinction between pied-piping of a full XP or a minimal operator (e.g., interrogatives/raising relatives vs. comparatives/complex adjectivals/non-raising relatives), a choice fixed by context (with various questions about relatives; see Sauerland 1998b). The wh-island analysis extends to other constructions if the feature that drives movement shares properties with wh- (assuming here a hierarchy of features);

TWO QUESTIONS: WHAT DO WE MEAN BY A HIERARCHY OF FEATURES? (WHERE DO WE CAPTURE THAT HIERARCHY? HOW IS THAT RELEVANT TO INTERVENTION?) AND SECOND: EXACTLY WHAT DO, SAY, TOPICALIZATION AND FOCALIZATION HAVE IN COMMON--THEY BOTH TRIGGER THE INTERVENTION EFFECT. WE CAN, OF COURSE, POSTULATE A COMPLETELY FORMAL FEATURE, BUT THEN IT IS JUST A WAY OF CODING THE MOVEMENT, NOTHING DEEPER. (I'M NOT CRITICIZING THAT, JUST GIVING IT A PROPER NAME.)

see Hagstrom (1998) for supporting evidence. Multiple overt wh-movement as in some Slavic languages might be analogous to multiple head options for A-movement along with a principle to overcome the subjacency effect.

PERHAPS, ALTHOUGH BOTH CLAIMS ARE HARD TO SEE. THOSE ARE NOT JUST MULTIPLE OPTIONS FOR MOVEMENT, BUT ACTUALLY MULTIPLE REAL MOVEMENTS. IN TURN, WHAT DOES IT MEAN TO OVERCOME AN EFFECT THAT FOLLOWS FROM SUCH A DEEP PRINCIPLE AS THE MLC (OR A SIMILAR VARIANT)? BESIDES, WHY DOESN'T OVERCOMING THIS EFFECT EXTEND TO A-MOVEMENT?

See Richards (1997), adapting ideas of Brody (1995); and for a general critical review, Boskovic (1998). As is well-known, pied-piping in A'-movement differs from the A-chain analogue, with variation among languages and constructions that is poorly understood.}

The reasoning

I ASSUME THIS IS THE BUSINESS OF PIED-PIPING ONLY THE HEAD...

extends to such constructions as (47) ((iii) based on Object Shift):

- (47) (i) *[John to seem [tI is intelligent]] (would be surprising)
- (ii) *(we hoped) [PRO to be decided [tI to be killed at dawn]]
- (iii) *[DO this book] seem [tDO to read [tDOI [never [[SU any students] t{read}]]]]

(iv) *there seem [a [SU several people]I are [PRED friends of yours]]

EPP is satisfied throughout, and if local matching sufficed for agreement, the expressions should converge with uninterpretable features deleted.

REMEMBER, THESE EXAMPLES ARE TRYING TO SHOW THAT, IN ADDITION TO WHAT GOES ON IN EPP TERMS, YOU NEED LONG-DISTANCE AGREEMENT TO TAKE CARE OF UNINTERPRETABLE FEATURES. THE QUESTION IS WHY ARE THE EXAMPLES IN (47) BAD IF THESE FEATURES COULD BE TAKEN CARE OF BY WAY OF EPP? JUST TO RECALL WHAT WE USED TO SAY (IF ANYTHING). IN (i), IF CHECKING John'S CASE STILL ALLOWS ITS CATEGORIAL FEATURE TO BE VISIBLE, WHY CAN'T John SUBSEQUENTLY MOVE TO THE SUBJECT OF seem, THEREBY SATISFYING THE EPP REQUIREMENT THERE? ABSTRACTLY, SOMETHING SIMILAR HAPPENS IN (ii). THE SUBJECT OF to be killed CAN RECEIVE NULL CASE, AS IN to be killed is no fun. SO PRO CHECKS ITS CASE FEATURE THERE, BUT STILL HAS ITS CATEGORIAL FEATURE AVAILABLE. WHY THEN CAN'T IT MOVE TO THE SUBJECT OF to be decided? TO MAKE THIS EXAMPLE RELEVANT, I ASSUME, WE MUST TAKE A STRUCTURE WHERE THERE IS NO NULL CASE IN to be decided, OR ELSE THE DERIVATION WOULD CRASH BECAUSE OF THAT. HOWEVER, THAT'S NOT IMPOSSIBLE IF THE TENSE WE CHOOSE IS OF THE DEFECTIVE TYPE. IN (iii), I HOPE, WE'RE NOT DEALING WITH ENGLISH. I IMAGINE THIS IS A LANGUAGE, FIRST OF ALL, WHERE THE VERB HAS RAISED OUT OF VP, MARKED BY never; I SUPPOSE THE DIRECT OBJECT IS TRYING TO MOVE OVER THE SUBJECT, VIA SOME INTERMEDIATE SPECS (vP AND TP, I IMAGINE), AFTER CHECKING ITS CASE IN THE vP SITE. THE MOVEMENT IS AS LOCAL AS THAT OF THE SUBJECT, ONCE THE VERB HAS RAISED TO T EXTENDING THE DOMAINS. OF COURSE, IN CHAPTER 4 WE WOULD HAVE SAID: SO WHAT? THAT DERIVATION CRASHES, SO TAKE THE OTHER ONE. THIS IS WHAT CHOMSKY WILL BE DENYING HERE, GOING INTO COMPLETE LOCALITY OF CHECKING FOR UNCLEAR REASONS. IN (iv) several people CHECKS ITS CASE DOWNSTAIRS AND ASSOCIATES AT LF TO there; IN CHAPTER 4 THIS WOULD NOT HAVE BEEN POSSIBLE BECAUSE there WAS ASSUMED NOT TO HAVE CASE, HENCE NOT TO BE ABLE TO CHECK THE CASE FEATURE OF seem. NOW CHOMSKY WILL TRY A UNIFIED ACCOUNT.

Appeal to such principles as "maximal checking" would not make the proper distinctions locally (e.g., barring (iii) in favor of subject raising).

I'M NOT SURE WHAT THIS "maximal checking" PRINCIPLE IS, ALTHOUGH JUDGING FROM THE FACT THAT (iii) IS MENTIONED, I SUPPOSE THE ISSUE IS CHECKING VIA OTHER ELEMENTS IN THE DERIVATION; FOR INSTANCE, IN (iii) THE QUESTION WOULD BE WHETHER, IF YOU MOVED AS INDICATED, THE SUBJECT WOULD CHECK ITS FEATURES. BUT SINCE CHOMSKY WANTS TO DISALLOW THIS SORT OF CALCULATION, WE DON'T HAVE AN ANALYSIS.

But in all cases the position superscripted "I" is inactive,

UNDER THE ASSUMPTION THAT CASE IS CHECKED IN THAT POSITION, INDUCING THE FREEZING EFFECT.

hence unable to raise ((i)-(iii)) or to delete the features of a matched probe (iv). Case (iv) illustrates a defective intervention effect of type (42): SU is visible (barring PRED as goal) but inactive, unable to establish agreement with matrix T.

THAT IS, several people IS INACTIVE BECAUSE IT IS CHECKING ITS CASE IN THE SPEC OF are, AND THUS IT CANNOT AGREE HIGHER UP IN THE TREE. NONETHELESS, IT ALLEDGELY PRODUCES AN INTERVENTION EFFECT, DISALLOWING friends of yours AS GOAL; OF COURSE, THE RESULTING SENTENCE WOULD BE A DISASTER (*there seem to friends of yours several people are), BUT PERHAPS FOR INDEPENDENT REASONS HAVING TO DO WITH THETA THEORY (friends of yours IS A PREDICATE).

Footnote {The perennial trouble-maker (i) falls into place if the (undeleted) [person] feature of embedded there bars association of matrix T to a man:

I SUPPOSE THIS IS A TYPO, AND CHOMSKY MEANS three men

lankspace1 line (i) there seem there to be three men in the room lankspace1 line

THIS IS LASNIK'S FAMOUS EXAMPLE. FOR THOSE OF YOU WHO ARE NOT NATIVE SPEAKERS, THE SENTENCE IS BAD (CHOMSKY SOMETIMES DOESN'T GIVE JUDGEMENTS). IN PRESENT TERMS, MATRIX T HAS TO AGREE WITH EMBEDDED three men, BUT IT CAN'T ACROSS THE [person] FEATURE OF EMBEDDED there, NOT DELETED IN THAT CONTEXT. I'M A BIT CONCERNED ABOUT WHY, IF A LESS THAN FULL AGREEMENT SET PREVENTS DELETION FROM OCCURRING, IT NONETHELESS INDUCES A BLOCKING EFFECT. THIS IS INTENTIONAL, AS FAR AS I CAN SEE, ALTHOUGH THE EVIDENCE ADUCED SO FAR (NAMELY (iv)) IS A BIT DUBIOUS). WHAT IS CLEAR, HOWEVER, IS THAT CHOMSKY WANTS TO SEPARATE THE INTERVENTION EFFECT FROM WHATEVER IS INVOLVED IN CHECKING, WHICH IS SENSITIVE TO FULLNESS OF FEATURE SETS FOR SOME REASON.

Groat (1997) points out further complications. Thus whatever its status, (ii) is more acceptable:

(ii) there look as though there are three men (vs. *[a man]) in the room

That suggests that the [person] barrier may be overridden in some manner.}

OR THAT (ii) IS THE ACTUAL PREDICTED INSTANCE, WITH LESS THAN FULL SETS *NOT* COUNTING FOR INTERVENTION EFFECTS. WHAT WOULD THEN BE THE PROBLEM WITH (i)? HOW ABOUT THIS: THE FEATURE OF THE INTERMEDIATE there HAS NOT BEEN CHECKED (DEFECTIVE T DOESN'T HAVE A FULL SET EITHER), AND THE DERIVATION CRASHES. IN OTHER WORDS, PERHAPS THE PROBLEM IS

HAVING there IN A DEFECTIVE T SITE, PER SE. OF COURSE, THE PROBLEM WITH THAT IS I believe there to be men there, WHICH IS GOOD. PERHAPS there IS, WHEN THERE IS A HIGHER v SITE, HIGHER UP IN THE TREE, A LA KOIZUMI (THAT v SITE WOULD NOT BE PRESENT IN LASNIK'S EXAMPLE, INVOLVING seem OR A MATRIX PASSIVE), AND HENCE PERSON MATCHING WOULD BE AGAINST v AND NOT AGAINST DEFECTIVE T.

The same property holds in (48):

(48)(i) *there were decided [a PRO to stay with friends] (ii) *XP T -seems that [a it was told friends CP]

PRO and it are inactive, their structural Case feature having been checked and deleted in a. But their f-features remain visible and block association of matrix T to friends, both of which therefore retain uninterpretable features.

PERHAPS I SHOULD SAY THAT THIS IS SLIGHTLY DIFFERENT FROM WHAT CONCERNED ME A MOMENT AGO. NEITHER pro NOR it INVOLVE LESS THAN FULL COMPLEMENTS OF FEATURES, SO THAT THEY SHOULD BE THERE FOR INTERVENTION EFFECTS SEEMS RATHER NATURAL (ANOTHER WAY OF SAYING THIS IS THAT WHAT CASE DOES TO YOU IS MORE THAN HAVE YOU THERE: IT GIVES YOU A SORT OF PASSPORT FOR TRANSFORMATIONS; INTERVENTION EFFECTS ARE NOT TRANSFORMATIONS, RATHER, THEY AFFECT THEM, AND FOR THAT YOU NEED NO PASSPORT). WHAT WORRIED ME WAS THE ISSUE OF WHETHER SOMETHING WHICH DOESN'T HAVE A FULL SET OF RELEVANT FEATURES, SUCH AS there OR ARGUABLY DEFECTIVE T, WHICH AFFECTS ITS VERY ACTIVITY AS AN OBJECT IN THE DERIVATION, FOR INSTANCE IN TERMS OF IDENTITY MATCHING FOR DELETION PURPOSES, SHOULDN'T BE ALSO AFFECTED IN ITS INTERVENTION PROPERTIES. THE ISSUE IS NOT ONLY EMPIRICAL; IT ALSO AFFECTS THE GENERAL DESIGN OF THE GRAMMAR.

Case (ii) is therefore barred with pure Merge of expletive, or raising of it or friends ("superraising").

Footnote{The MP analysis of superraising assumed that it raises to matrix T by locality and the derivation crashes by lack of Case-checking for friends. But covert attraction of the f-features of friends by matrix T should still be possible, with convergence (Eduardo Raposo, pc). See Raposo and Uriagereka (1996) for a different approach.}

OUR APPROACH PREVENTED 'DISPARATE CHECKING', WHERE A GIVEN HEAD CHECKS, SAY, EPP, AND AN ENTIRELY DIFFERENT HEAD CHECKS CASE, AND THOSE TWO DO NOT FORM A UNIFIED OBJECT (AS IN WEAK EXPLETIVES).

We are now in a position to derive the basic structural properties of CFCs ((5), extended to (50)), in the configuration (49):

(49) $a = [XP [(EA) H YP]]$

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

(ii) In the configuration $[\beta H\beta \dots a]$, $H\beta$ a CFC and β minimal,

(a) if $H\alpha$ is C, $H\beta$ is independent of a

(b) if $H\alpha$ is v, $H\beta = T\beta$ agrees with EA, which may raise to SPEC- $T\beta$ though XP cannot

(c) if $H\alpha$ is Tdef, if $H\beta$ is T then XP raises to SPEC- $T\beta$ if there is no closer candidate G for raising; and if $H\beta$ is v then XP agrees with v (as may a lower associate if $XP = EXPL$).

(50i) has already been dealt with (pp. 00, 00), so we can keep to (ii).

In case (a), $a = [XP [C TP]]$, TP headed by nondefective $T\alpha$. We can limit attention to $T\alpha$ with its f-set deleted; otherwise the derivation would have crashed at a.

RECALL, THIS IS EXPRESSED WITHIN A COMPLETELY LOCAL SYSTEM, SO IF YOU DON'T TAKE CARE OF BUSINESS AT a (E.G. DELETING ITS UNINTERPRETABLE FEATURES), YOU FINISH THEN AND THERE.

Hence agreement is fully established with the closest associate ASSOC that matched $T\alpha$; ASSOC is either raised to SPEC- $T\alpha$ or remains in situ in a long-distance agreement relation, its structural Case feature deleted in either case. Any structural Case feature remaining in a is inaccessible by virtue of the defective intervention effect induced by ASSOC, again causing crash detectable at a.

THE ONLY CAVEAT I ADDED HERE IS WHAT HAPPENS WHEN THE ASSOCIATE DOESN'T HAVE A FULL COMPLEMENT OF FEATURES (there). AS WE SAW, THE DATA ARE NOT COMPLETELY CLEAR (GROAT'S EXAMPLE IS THE KEY), AND THE CONCEPTUAL ISSUES ARE OPEN.

Case (a) follows: the f-set and EPP-feature of $H\beta$ have to be satisfied independently of a.

Consider case (b), with $a = [XP [EA [v YP]]]$. XP is raised from within YP, checking and deleting the f-set of v (object agreement) and its own structural Case feature. XP is therefore inactive for A-movement, and cannot check the f-features of $T\beta$. These can only be deleted by agreement with EA, deleting its structural Case feature as well, with EA either raising to SPEC- $T\beta$

THAT WOULD BE THE FAMILIAR CASE OF SUBJECT MOVEMENT

or remaining in situ.

THOSE ARE THE ICELANDIC CASES.

Recall that XP does not induce a defective intervention effect that would bar the TB-EA relation; see (41).

THAT'S BECAUSE OF EQUIDISTANCE.

Footnote {We are, furthermore, speculating that XP can only appear if forced

XP IS NOT THERE BY MERGE, BUT BY MOTIVATED MOVEMENT, THUS:

(barring EXPL-raising always,

THAT IS, VIA v , WHICH IS WHAT WE'RE DISCUSSING NOW. BUT I'M LOST HERE. I TALKED TO RAPOSO ABOUT THIS, AND HE'S CONFUSED TOO, AND WE CONSTRUCTED THE FOLLOWING EXAMPLE OVER THE PHONE: there seem to me t to be a proof discovered. THE QUESTION IS HOW THE RAISING OF there CAN PROCEED IF IT ISN'T THROUGH THE v OF seem, GIVEN THE IMPENETRABILITY CONDITION. UNLESS CHOMSKY WANTS TO SAY THAT THE v OF seem IS SOMEHOW DIFFERENT, AND IT DOESN'T INVOKE A PHASE. BUT THIS NEEDS, I BELIEVE, SOME SERIOUS CLARIFICATION.

and argument raising except to implement IFM

THAT IS, BASICALLY, SUCCESSIVE CYCLICITY

or LF-interpretive effects);

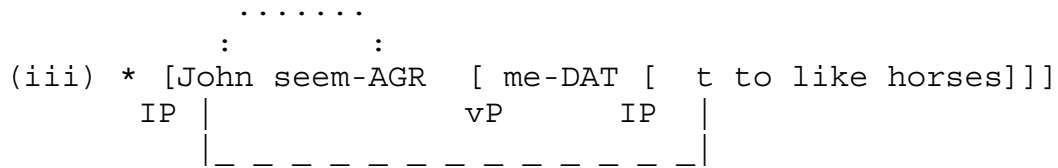
WHATEVER THAT MEANS, AS WE SAW.

see discussion of (24), (25). I overlook questions of possible parametrization for OVS languages; see Ura (1996).}BAR

In case (c), $a = [XP [T_{def} YP]]$. Whether merged (hence EXPL) or raised, XP has an uninterpretable feature that renders it active for the Case-agreement system, and must be checked outside a by $H\beta$ which is either v (ECM) or T (raising). Suppose $H\beta$ is v . If $XP = EXPL$, its uninterpretable feature [person] deletes but the f -set of v remains, able to check the Case (accusative) of a lower associate ("I expect [there to be a proof discovered]"); there is no defective intervention effect.

I ASSUME THIS IS BECAUSE THE FEATURE OF THE EXPLETIVE IS LITERALLY GONE, UNLIKE INTERPRETABLE FEATURES THAT REMAIN EVEN AFTER RENDERED INACTIVE FOR CHECKING.

If XP is an argument, it is the associate of v and object-agreement with (accusative) Case-checking proceeds as before. In either case, XP raises to SPEC- v only when v has an



(i) IS A BIT LIKE THE OLD ENGLISH Methought they..., EXCEPT SHOWING AGREEMENT WITH THE SUBJECT OF THE EMBEDDED CLAUSE, AND (iii) IS A BIT LIKE THE MODERN ENGLISH John seems to me to ..., EXCEPT THE SENTENCE IS BAD IN ICELANDIC. AS FOR (ii), IT IS SIMILAR TO (i), EXCEPT AGREEMENT IS DEFAULT, AND OF COURSE THE CLAUSE ITSELF IS DIFFERENT IN SHAPE.

In (i), the matrix verb agrees with the embedded nominative, but not in (ii), which requires default inflection because of the defective intervention effect: the f-features of John block the T-associate relation between T-seem and nominative horses.

RECALL ALSO THAT John ITSELF CANNOT AGREE, SINCE IT HAS INHERENT CASE, AND THAT REGARDLESS OF WHAT ITS CASE STATUS MIGHT BE, IT INDUCES A BLOCKING EFFECT.

In (i) as well a phrase with f-features intervenes between matrix T and NOM, namely the trace of the quirky dative me. But the latter is not the head of an A-chain, in contrast to (ii) (also (47iv) and (48)).

SO IT IS NOT JUST POSITIONS THAT BLOCK RELATIONS (IN THIS INSTANCE A TRACE), BUT ACTUALLY CHAINS--AS IT WAS IN CHAPTER 4.

Case (iii) is blocked by locality; quirky DAT, with structural Case, is accessible.

THAT IS ONE OF THE REASONS YOU MUST ASSUME, IN THIS SYSTEM, THAT QUIRKY IS STRUCTURAL, IN ADDITION TO BEING INHERENT (THUS NOT INDUCING AGREEMENT). OF COURSE THE QUESTION IS WHY (iii) IS GOOD IN ENGLISH, BUT WE'VE TALKED ABOUT THAT ALREADY.

The conclusions are as before, but sharpened: it is only the head of the A-chain that blocks matching under the locality condition (40iii). A-movement traces are "invisible" to the probe-associate relation; or from another perspective, the A-chain itself (regarded as a set of occurrences of a) constitutes the barrier.

This account relies on restriction of basic operations to Merge and Agree, based on feature-matching (identity) and driven by Suicidal Greed. Suppose that there is a Move-a operation dissociated from matching and dependent on properties of a, with locality expressed as "shortest Move." In the defective intervention cases, matrix T is the closest target that can check the Case feature of the potential associate a. With Move-a available and constrained only by locality, the derivation converges (incorrectly) with raising or agreement of a. Further constraints are required, then, if this option exists.

FRANKLY, THIS PROBLEM IS TOO TECHNICAL TO MERIT, IN ITSELF, THE REDESIGN OF THE GRAMMAR. IN FACT, FOOTNOTE 50 OF CHAPTER 4 HAD AN APPROPRIATELY TECHNICAL SOLUTION: ONE WAY TO PROCEED WAS 'to interpret overt erasure of F as meaning conversion of F to phonological properties, hence stripped away at Spell-Out'. IF SO, ONE NEED NOT TOUCH SPELL-OUT. OF COURSE, I BELIEVE SPELL-OUT SHOULD BE TOUCHED, BOTH FOR CONCEPTUAL AND (OTHER) EMPIRICAL REASONS, BUT I WOULDN'T BANK THE QUESTION ON THE ERASURE OF FEATURES.

The natural conclusion is that Spell-Out is associated with agreement. Deleted features are literally erased, but only after they are sent to the phonological component along with the rest of the structure S -- possibly at the phase level. Spell-Out therefore applies cyclically in the course of the (narrow syntactic) derivation. I will assume that this approach, apparently the simplest and most principled one, is correct.

Footnote{The conception is similar to the strict derivational interpretation of Spell-Out of Epstein et al. (1998);

AS FAR AS I CAN SEE, IT IS LITERALLY THE SAME, AND INDEED IT GOES BACK TO WORK ON THESE MATTERS IN THE EARLY SEVENTIES, BY BRESNAN, JACKENDOFF, AND LASNIK (SEE BELOW, WHERE BRESNAN IS MENTIONED).

see Uriagereka (1996, forthcoming b) for similar suggestions, on different grounds.

DIFFERENT *EMPIRICAL* GROUNDS, BUT EXACTLY THE SAME KIND OF ARCHITECTURE. AN ISSUE WILL REMAIN, THOUGH, ABOUT WHAT IS THE DOMAIN OF SPELL-OUT, WHETHER PHASES, AS SUGGESTED BY CHOMSKY (WHAT I CALLED FLOWS), OR RATHER THAT *AND* WHAT I CALLED CASCADES, THE ESSENCE OF CED EFFECTS AND 'SEPARATE' MERGES.

The basic architecture resembles that of Bresnan (1971); her results on the interaction of phonological and transformational rules fall into place more directly than in the EST model. See several papers in Abraham et al. (1996) and Yang (1997) for related discussion.}

The single Spell-Out thesis of MP retains the flavor of the EST model, distinguishing overt from covert operations -- pre- and post-Spell-Out, respectively. If both overt and covert operations are cyclic, then there are two independent cycles; and if operations of the phonological component are cyclic, a third cycle as well. With cyclic Spell-Out, contingent on feature-checking operations, these distinctions collapse. There is a single cycle; all operations are cyclic. Within narrow syntax, operations that have or lack phonetic effects are interspersed. There is no distinct LF component within narrow syntax, and we can dispense with troublesome questions about its apparently cyclic character.

UNDERSTAND THIS WELL: IT'S NOT THAT CHOMSKY IS DENYING LF, BUT RATHER THE NEED FOR A *SEPARATE* LF COMPONENT, WHERE VARIOUS SPECIFIC OPERATIONS (QR, ETC.) CAN PROCEED. EMPIRICALLY, THOUGH, THE

OBJECT YOU OBTAIN AT THE END OF THE DERIVATION IS THE SAME WE HAD IN CHAPTER 4, IN FACT IDENTICALLY SO IF YOU KEEP A *LEVEL* OF LF (AS CHOMSKY DOES), WHICH FORCES A FINAL UNIFICATION.

Agree alone, not combined with Merge in the operation Move, can precede overt operations, contrary to the assumptions of MP and related work.

THIS IS CONCEPTUALLY DIFFERENT FROM CHAPTER 4, AND AS WE SAW ALSO EMPIRICALLY DIFFERENT.

Crucial cases include long-distance agreement, wh- in situ, and others. Many questions arise, but they do not seem obviously unanswerable.

One question has to do with identification of chains by identity under cyclic Spell-Out. Suppose the uninterpretable features of a head H are checked at phase PH with PH transferred to the phonological component and H reduced to H', which is transferred at a later phase of derivation. We want to take H', H to be a chain, so that H will be unpronounced by general principles, but H, H' are not identical. We may therefore take chains to be determined by a relation of non-distinctness holding between (a, β) if they differ only in uninterpretable features.

THIS LOOKS LIKE A VERY TECHNICAL SOLUTION TO THE VERY SERIOUS QUESTION OF IDENTITY. N&U BITE THIS BULLET IN A DIFFERENT DIRECTION, IN FACT FORCING THE SPELL-OUT OF HEADS AND AND TAILS OF CHAINS TO ENSURE IDENTITY, BUT WITHIN A SET OF DIFFERENT ASSUMPTIONS (WE'LL RETURN TO THAT PAPER). I THINK THE ISSUE IS REAL, EVEN IF:

Footnote {The issue is narrow, and might be avoidable in other ways.

NOT COMPLETELY CLEAR TO ME HOW, SINCE I THINK IT GENERALIZES WHEN YOU LOOK AT MORE COMPLEX CHAINS.

It would be restricted to subjects in Object Shift constructions if the "reconstruction" operation induced by the final step of A'-movement preempts deletion, so that strict identity holds throughout A'-chains.}

I GUESS WHAT THIS MEANS IS THAT... NO, I DON'T KNOW WHAT THIS MEANS. JUST TOO CRYPTIC.

The principle Procrastinate is no longer formulable (at least, as before), eliminating another case of look-ahead.

BASICALLY WE DON'T NEED TO SPEAK IN TERMS OF PROCRASTINATE, SINCE THERE IS NO SIGNIFICANT LF COMPONENT, JUST THE LF REPRESENTATIONS.

The concept strength, introduced to force violation of Procrastinate, appears to have no place. It remains to determine whether the effects can be fully captured in minimalist terms or remain as

true imperfections.

To implement the program in a suitably Spartan fashion

I LOVE THAT PHRASE! I SUPPOSE WE COULD USE IT WHEN REFERRING TO THE STRONG OR ONTOLOGICAL THESIS...

we seek to restrict CHL to indispensable operations that satisfy minimalist conditions. There are two candidates: Merge, indispensable in some form, and Agree, which, we speculate, might ultimately be accounted for in terms of design specifications. We have to determine how these operations apply, seeking and questioning departures from optimal design.

The operations Merge and Agree must:

- (52) (I) Find syntactic objects to which they apply
- (II) Find a feature F that drives the operation
- (III) Perform the operation, constructing a new object K.

An operation OP takes objects already constructed (perhaps in the lexicon),

WOW, MAJOR WORDS! *IN* THE LEXICON? OR IN THE NUMERATION? WHAT EVER HAPPENED WITH THE 'REPOSITORY OF IDIOSYNCRACIES'? WE WILL NEED THAT, CALL IT LEXICON OR WHATEVER. AT THE SAME TIME, I THINK WE WILL NEED SOME CONSTRUCTIVE SPACE FOR LEXICAL STUFF, CALL IT LEXICON OR WHATEVER. PERSONALLY (AND THIS IS JUST NOMENCLATURE THAT I'M PERFECTLY WILLING TO CHANGE), I'D CALL 'LEXICON' THE LIST OF IDIOSYNCRACIES, AND WOULD ASSEMBLE STUFF, IN AS CREATIVE A WAY AS I NEED TO IN THE NUMERATION. THAT STILL DOESN'T MAKE THE NUMERATION A SYNTACTIC LEVEL OF REPRESENTATION.

and forms from them a new object. Condition (I) is optimally satisfied if OP applies to full syntactic objects already constructed, with no search; that is, if CHL operates cyclically. It follows that derivations meet the condition (53):

- (53) Properties of the probe/selector a must be exhausted before new elements of the lexical subarray are accessed to drive further operations

I'M NOT SURE I COMPLETELY FOLLOW THIS, BUT I SUPPOSE IT FORCES A VERY LOCAL ACTIVATION. MY CONCERN IS THIS. TAKE THE NUMERATION {the, man, saw, a, woman} AND SUPPOSE YOU START OPERATING WITH saw. YOU'RE ALREADY IN TROUBLE, BECAUSE NOTHING YOU CAN PULL FROM THE NUMERATION DOES THE TRICK. THE NUMERATION WOULD WORK IF YOU HAD it, OR SOMETHING. BUT ALL YOU HAVE IS THE *POSSIBILITY* TO FORM a woman, FOR INSTANCE, WHICH MEANS ACCESSING AN ELEMENT FROM THE NUMERATION PRIOR TO SATISFYING THE PROPERTIES OF saw. OF COURSE, YOU COULD SAY THE PROBLEM WOULD NOT ARISE IF YOU STARTED THE DERIVATION WITH a, WHICH

CAN YIELD, a woman, AND THEN YOU GET saw. BUT NOW WHAT? BACK TO WHERE YOU STARTED: YOU NEED TO GET the AND MERGE IT PRIOR TO saw. AND SO ON. WHICH AS FAR AS I CAN SEE MEANS THAT WITHIN A CERTAIN DOMAIN (CALL IT A PHASE IF YOU WISH, IF THAT'S THE RELEVANT DOMAIN), BASICALLY ALL BETS ARE OFF, AND YOU CAN DO WHATEVER YOU WANT IN TERMS OF APPLYING OPERATIONS. I'M NOT SAYING THIS IN A CONVINCED WAY, BUT POSING IT AS A TOPIC OF DISCUSSION FOR THIS COURSE.

If the properties of a are not exhausted, the derivation crashes because a can no longer be accessed.

OF COURSE I'M WILLING TO BUY THAT *ACROSS PHASES*. IN FACT, THAT WAS THE WHOLE INTUITION BEHIND THE MULTIPLE SPELL-OUT SYSTEM, ALTHOUGH THERE THE ISSUE HAD NOTHING TO DO WITH ECONOMY IN THE DESIGN OF OPERATIONS, BUT WAS TAKEN TO BE A MERE MATTER OF WHETHER THE DERIVATION IS ACTIVE (PRIOR TO SPELL-OUT) OR NOT (AFTER SPELL-OUT). AGAIN, THIS IS A BIG TOPIC THAT I'D LIKE TO CONSIDER HERE.

We continue to assume cyclic application of all operations, returning to some ambiguities.

Let's turn now to condition (III).

The operation Merge forms K from a, β . Minimally, K should consist only of a and β , so $K = \{a, \beta\}$. More information is needed about K, however: its category (its label) and the nature of the merger, either substitution or adjunction -- the former at least not entering into narrow syntax, on the sparsest assumptions, but needed for the phonological component and LF-interpretation.

DON'T KNOW WHAT THAT MEANS, FOR SUBSTITUTION...

Footnote {The discussion of categories and labels here largely follows Collins (1997). Questions about categories raised in note MAR are not relevant here;

THAT WAS FOOTNOTE 86, ON CATEGORIAL FEATURES. THAT'S DIFFERENT FROM WHETHER THE OBJECT HAS A LABEL; IT MAY HAVE NO CATEGORIAL FEATURES, AND YET BE LABELED:

the syntactic object has some distinctive property that must be recognized.}

ACTUALLY, WHETHER IT MUST, AND WHEN IT MUST BE RECOGNIZED IS THE CENTRAL TOPIC OF H&U, WHICH I RETURN TO.

To ensure that every category has a label, let us say that $\text{label}(a) = a$, for an LI.

UNFORTUNATELY THIS SIDESTEPS THE ISSUE POSED BEFORE, POSSIBLY IN A DANGEROUS WAY, SINCE IT IS NOT OBVIOUS THAT THE LABEL FOR AN OBJECT SHOULD BE THE OBJECT, JUST AS IT IS NOT OBVIOUS THAT AN OBJECT IS

IDENTICAL TO THE SET INCLUDING THAT OBJECT (OR ELSE, SAY, THE EMPTY SET SHOULD BE IDENTICAL TO THE SET CONTAINING THE EMPTY SET, BUT THAT'S CONTRADICTORY). I HONESTLY BELIEVE THE MATTER IS NOT JUST NOTATIONAL, BUT WON'T GO INTO THIS NOW.

Adjunction has an inherent asymmetry: X is adjoined to Y. Exploiting that property, let us take the distinction between substitution and adjunction to be the (minimal) distinction between the set $\{a,\beta\}$ and the ordered pair a,β , a adjoined to β .

I MUST STEP HERE AGAIN WITH ANOTHER FORMAL POINT, THE ORDERED PAIR a,β IS SOMETHING LIKE THE SET $\{\{a\}, \{a, \beta\}\}$ (THE ANGLED BRACKETS ARE JUST A NOTATIONAL TRICK). I FIND THIS SOMEWHAT INTRIGUING, BECAUSE BY THE MERE FACT THAT YOU HAVE TO LINEARIZE $\{a,\beta\}$ INTO $\{\{a\}, \{a, \beta\}\}$ WHEN YOU PRONOUNCE IT, YOU ARE CREATING A TERM $\{a\}$, WHICH YOU CAN DISTINGUISH FROM THE LABEL a IN TRIVIAL WAYS. ANOTHER WAY OF SAYING THIS IS AS FOLLOWS: PERHAPS LINEARIZATION HAS AN INTERESTING SIDE-EFFECT: IT OBJECTUALIZES LABELS, IN A DIRECT WAY. SIMILARLY FOR ADJUNCTION. PERHAPS BY ADJOINING a TO β AND THUS OBTAINING $\{\{a\}, \{a, \beta\}\}$ YOU ARE IN EFFECT OBJECTUALIZING a WITH NO CODING TRICKS. I'M ASSUMING IN THIS LITTLE DISCUSSION THE DEFINITION OF TERM IN CHAPTER 4, AND A NATURAL DEFINITION OF LABEL (A TERMINAL WHICH IS NOT A TERM). NOTE THAT THE BASE OF THAT DEFINITION HAS TOP OBJECTS (INCLUDING WORDS SELECTED FROM THE NUMERATION) AS TERMS, THE BASE STEP. THAT MEANS YOU WOULD BE OPERATING WITH UNLABELED TERMS. INDEED, LABELS WOULD ONLY ARISE AS YOU LINEARIZE THE OBJECTS, THUS AT SPELL-OUT. THE QUESTION IS WHETHER YOU NEED LABELS PRIOR TO THAT. IF YOU DO, WHAT I SAID CAN'T BE RIGHT.

The constructed objects K, then, are of the form $\{G, \{a,\beta\}\}$ (substitution) or $\{G, a,\beta\}$ (adjunction), G the label of K.

The term "substitution," adapted from earlier work, is misleading within this framework. For clarity, let us refer to substitution as set-Merge and adjunction as pair-Merge. I will put aside here a number of questions that arise about adjunction.

WE'LL RETURN TO SOME OF THEM IN M&U, FOLLOWING IDEAS THAT GO BACK TO CHAMETZKY.

Footnote{See Kayne (1994), MP. Kayne assimilates adjunct and specifier. MP suggests that XP-adjunction may not be part of narrow syntax (chap. 4.7.3); possibly not a central range of cases of head-adjunction either (see notes TYP, HADJ).}

On minimal assumptions, the label G should be the label of either a or β . Hence no matter how complex the object constructed, its label is an LI, the head selected from the lexicon that has "projected" through the derivation, or a reduced MLI. If the label is determined from a,β by general principles, then the result of merger of a,β is simply $\{a,\beta\}$ or a,β .

NOTE, THAT WAS AN 'IF'; SO:

Are labels predictable?

Consider pure Merge. There are two cases: set-Merge and pair-Merge. The latter adjoins a to β to form a,β . Given the asymmetry, it is natural to conclude that the adjoined element a leaves the category type unchanged: the target β projects. Hence adjunction of a to β forms $K = \{G,a,\beta\}$ where G is the label of β . Eliminating redundancy, the operation forms $K = a,\beta$.

KEEP THIS CONCLUSION IN MIND, SINCE IT WILL BE RATHER IMPORTANT WHEN WE READ H&U.

As an operation, set-Merge is symmetric, so one might expect either label to project. If so, the outcome would either be interpretable at LF or not. But here too properties of language design appear to determine the label without look-ahead. Set-Merge typically has an inherent asymmetry. When a,β merge, it is to satisfy (selectional) requirements of one (the selector) but not both. Fairly generally, furthermore, the selector is uniquely determined for a pair (a,β) , as can be seen by reviewing cases.

Set-Merge of (a,β) has some of the properties of Agree: a feature F of one of the merged elements (say a) must be satisfied for the operation to take place.

IN EARLIER VERSIONS OF THE PAPER, WHEN AGREE WAS STILL NOT INTRODUCED, THIS SENTENCE LIKENED SET-MERGE TO ATTRACT. THE POINT IS HE'S GOING TO SAY THERE'S SOMETHING LIKE A PROBE AND A GOAL. WE'LL SEE THAT THERE MAY BE NON-TRIVIAL INSTANCES OF THIS IN RELATIONS BETWEEN ADJUNCTS AND HEADS, WHEN WE READ M&U.

Furthermore, F is in the label of a , hence detectable in an optimal way (satisfying condition (52II)). The selector F for Merge is analogous to the probe for Agree. F is, furthermore, the only element of a that enters into the operation, hence the only one available without further complication to determine the label of the merged elements. In this case too, then, the label is predictable and need not be indicated: the label of the selector projects.

NOTICE, INCIDENTALLY, THAT THIS GIVES THETA-THEORY A PRIMITIVE STATUS. IN OTHER WORDS, THE APPROPRIATE RELATION BETWEEN THE MERGED ITEMS IS NOT JUST WHATEVER SEMANTICS TELLS YOU LATER ON. RATHER THERE IS A SYNTACTIC PROCESS AKIN TO AGREEMENT, OR PERHAPS WE SHOULD SAY THERE IS A GENERALIZED SYNTACTIC OPERATION INVOLVED IN BOTH AGREE AND THIS KIND OF SELECTION, CALL IT SYN, WHICH INVOLVES PROBES, GOALS, ETC.

The intuitive content of the Projection Principle is that for a substantive category a , the selector F is a semantic property of G , an interpretable feature. Hence F does not delete (another difference between Theta and Checking Theories).

ANOTHER INDICATION THAT THETA THEORY IS REAL; APART FROM BEING EXECUTED UNDER MERGE, IT DOESN'T INVOLVE F DELETION.

Further specification depends on how Theta Theory is understood. Take, say, transitivity of a verb V. If the property is implemented in terms of Q-grids, then a feature of V selects the object. If it is implemented configurationally as a structure [v v [V V...]], then v too is a relevant selector and the v-VP structure enters into the interpretation. In either case, one or another form of deviance (or crash) results from inappropriate merger. These and many other questions come to the fore as the framework is more carefully articulated.

The two cases of pure Merge differ in several respects. The asymmetrical operation pair-Merge has no selector and is optional; the symmetrical operation set-Merge has a selector (typically unique) and is obligatory.

SUGGESTING THAT PAIR-MERGE DEFINES ITS OWN COMPONENT, DIFFERENT FROM THETA THEORY AND FROM WHATEVER GOES ON IN STANDARD CHECKING, THUS SUGGESTING ADJUNCTS ARE IN A DIFFERENT DIMENSION. THIS MUST SURELY RELATE TO THEIR PECULIAR EXTRACTION PROPERTIES AND OTHERS, A MATTER WE RETURN TO WHEN WE READ M&U.

Footnote {On the conventional and reasonable assumption that the selectional feature itself may be optional, as in the case of a verb that may or may not have IO ("send," etc.).}

THERE'S MORE THAN MEETS THE EYE TO THIS FOOTNOTE, BUT LET'S RETURN TO THIS WHEN WE REVIEW THE MORI'S WORK LATER ON IN THE SEMESTER.

In these respects language design is close to optimal, providing just the information necessary for an operation OP to project the label L: L is determined by OP itself if OP is asymmetrical, but a selector is needed to determine L if OP is symmetrical.

OBSERVE: THE CLAIM IS THAT THE REASON YOU NEED A SELECTOR IS THAT MERGE ITSELF IS SYMMETRICAL. IF THIS IS CORRECT, THETA THEORY IS A REQUIREMENT IN THE SYSTEM, JUST AS, SAY, A LINEARIZATION PROCEDURE IS (BE IT LCA OR ANY OTHER). IN PARTICULAR, GLOBALITY CONSIDERATIONS OF THE SORT DISCUSSED THROUGHOUT DEMAND THETA THEORY.

Accordingly, Merge has a selector for set-Merge but not pair-Merge, which is therefore optional. The label is determined without look-ahead to check eventual convergence (contrary to MP), another case confirming conclusions about complexity suggested earlier.

Pure Merge, then, satisfies the conditions of (52) rather well. Consider the second elementary operation, Agree. We are now assuming cyclicity: a is a candidate for the operation only

I GUESS only IF

it is the full syntactic object under inspection, so search is unnecessary, satisfying (I).

RECALL THAT CHOMSKY STIPULATES CYCLICITY IN ORDER TO MAKE SEARCH UNNECESSARY (NOT SOMETHING AS DRASTIC AS NOT HAVING THE STRUCTURE AVAILABLE, AS IN MSO).

Agree requires a probe F in a. By condition (II), F has to be readily detectable, hence optimally in the label L(a) of a, its sole designated element.

F IS, AS IT WERE, IN THE 'TIP OF THE ICEBERG', SEEKING F' INSIDE THE ICEBERG ITSELF.

F seeks a matching feature F'. Deletion takes place under conditions already discussed. The syntactic object a is otherwise unchanged. Its label remains as before. The same properties carry over to Move, constructed from Agree and Merge.

Footnote{An elaborate argument to guarantee projection of the target in MP is superfluous, under this reanalysis. The label is determined by the probe for the Agree component of Move, and by the selector for the Merge component. These are different features, but they yield the same choice of label.

I GUESS I DON'T KNOW ENOUGH ABOUT THE STRUCTURE OF LABELS TO DETERMINE WHETHER THIS IS NECESSARY, ALTHOUGH IT IS PLAUSIBLE.

It seems a notational question whether we take the label to be the original LI or a reduced MLI.}

In all cases, then, the label is redundant. The syntactic objects are LIs, or sets {a,β} or a,β constructed from them. The label is determined and available for operations within CHL or for interpretation at the interface, but is indicated only for convenience.

ACTUALLY, IT IS JUST FOR CONVENIENCE THAT THE *INDICATION* TAKES PLACE, BUT THE LABEL IS QUITE REAL; THE FACT THAT ITS PRESENCE FOLLOWS FROM THE INTERACTION OF SOMETHING ELSE DOESN'T MEAN THAT IT IS NOT A REAL ELEMENT IN THE SYSTEM. IN FACT, YOU'LL SEE WHEN WE READ H&U THAT WITHOUT THAT ASSUMPTION THEIR ENTIRE SYSTEM COLLAPSES, AND WITH IT ALL THE EMPIRICAL PREDICTIONS THEY MAKE.

Computation is driven by a probe/selector of a label, which projects. Hence no operation can be contingent on application of earlier ones. That seems to be a valid descriptive generalization, which falls into place. Both label-determination and operations are "first order Markovian," requiring no information about earlier stages of derivation.

AGAIN, THIS INVOLVES ABSOLUTE LOCALITY. WE WILL QUESTION LATER ON WHETHER SUCH A DRASTIC LOCALITY IS DESIRABLE, GIVEN THE FACTS.

Conditions (52I,II) are optimally satisfied, as is (III) for Merge and partially (so far) for Agree.

Continuing with Agree, (III) requires that the matched goal G must be easily located. We want to identify a domain D(P) of the probe P, such that G is within D(P). There are two candidates for D(P): the smallest and the largest of the categories labelled by the label containing P.

IT IS INTERESTING THAT CHOMSKY INVOLVES smallest AND largest CATEGORIES. SUCH NOTIONS ARE STANDARD IN THE SCIENCES OF COMPLEXITY, INVOLVING NATURAL OBJECTS (OFTEN AS 'MAXIMUMS' AND 'MINIMUMS'). NOTICE THAT FOR THOSE EXTREMES YOU DON'T NEED ARTIFICIAL LABELINGS OR CODING SUB-PARTS IN TERMS OF AN ORDERING (E.G. NUMBERED LAYERS, ETC.). IT'S A VERY ELEGANT WAY OF IDENTIFYING NATURAL DOMAINS.

The former includes only the complement of P; the latter its specifiers as well. Search space is more limited if D(P) is the smallest category, as so far assumed; see (40II).

SAME GAME AS BEFORE: REDUCE COMPLEXITY AS MUCH AS POSSIBLE. OBVIOUSLY, THIS IS A CLAIM ABOUT THE INTERNAL WORKINGS OF THE SYSTEM, NOT OUR THEORY. THEORY-WISE IT WOULD HAVE BEEN EQUALLY COMPLEX TO GO WITH THE LARGEST CATEGORY.

It is restricted further by the "closest match" condition (40III)). Natural complexity/economy conditions are again satisfied, along the lines of (3), (14), and their extensions; there is substantial evidence that G must be in the complement of the probe P, not its specifiers, and that locality conditions enter into choice of G.

ALTHOUGH AS I SAID BEFORE NOT JUST THE COMPLEMENT DOMAIN, BUT COMPLETELY WITHIN A 'COMPLEMENT PATH' (NEVER INSIDE A SPECIFIER). THE MSO SYSTEM PREDICTS THAT, AND IS AS FAR AS I CAN SEE COMPATIBLE WITH EVERYTHING ELSE BEING SAID SO FAR. AT ANY RATE, WHETHER IT IS WITH THAT SYSTEM OR WITH SOMETHING ELSE, YOU MUST MAKE SURE THAT YOU DON'T SEEK YOUR GOAL OTHER THAN IN A 'COMPLEMENT PATH'.

The basic operations Merge and Agree satisfy reasonable "good design" conditions (52). The conclusions extend to the third operation, Move, insofar as it is constructed from the basic operations.

Move of β , targeting a, has 3 components:

(54)(I) A probe P in the label L of a locates the closest matching G in its domain

(II) A feature G' of the label containing G selects a phrase β as a candidate for "pied-piping"

(III) β is merged to a category K

P and G' are uninterpretable. P deletes if G is active (Suicidal Greed). G' also deletes, but it cannot delete in step (I) before carrying out its function in step (II). There are reasons to suppose that G cannot delete before (III), but I defer the matter.

I'M NOT SURE I KNOW WHAT THOSE REASONS ARE. I DON'T THINK THIS IS WHAT CHOMSKY HAS IN MIND, BUT IN THE BASQUE PAPER WE'LL CONSIDER LATER ON, I HAVE TO SAY THAT G DELETES AFTER (III), WHICH IS WHAT CREATES A SPEC-INDUCED ISLAND. WE'LL RETURN TO THAT, BUT DOES ANYONE KNOW WHAT CHOMSKY MEANS HERE, AND WHY HE'S DEFERRING IT?

The new object K formed by merge of β to a retains the label L of a, which projects. There are two reasonable possibilities, illustrating the ambiguity of cyclicity mentioned earlier:

- (55)(a) a is unchanged
- (b) β is as close to L as possible

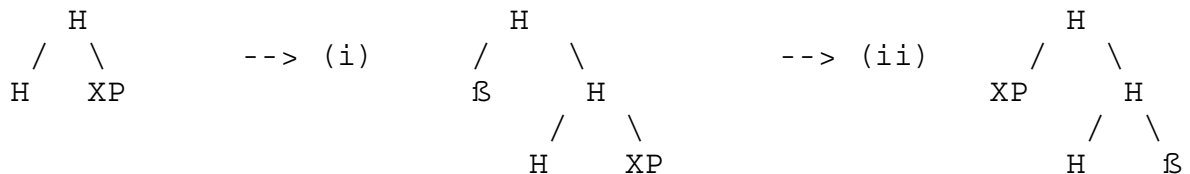
Suppose we have the LI H with selectional feature F, and XP satisfying F. Then first Merge yields $a = \{XP, H\}$, with label H. Suppose we proceed to second Merge, merging β to a. In this case, β is either extracted from XP (Move) or is a distinct syntactic object (pure Merge). There are two possible outcomes, depending on choice of K in (55):

WHAT HAPPENS IN SECOND MERGE IS HARD TO VISUALIZE. I DON'T KNOW WHETHER THIS WILL COME OUT RIGHT AFTER TRANSLATING THE PAPER TO HTML, BUT IF NOT FOLLOW THE IMAGE IN THE WHITEBOARD. IN ANY CASE, I FIND THESE IDEAS EASIER TO UNDERSTAND WHEN SEEING THE TREES IN (57).

- (56)(i) $\{\beta, \{XP, H\}\}$ (as in (55a))
- (ii) $\{XP, \{\beta, H\}\}$ (as in (55b))

In tree notation, the equivalents are (57i,ii), respectively:

(57)



THAT IS IN (i) YOU MERGE AT THE ROOT, WHEREAS IN (ii) YOU MERGE AT THE FOOT.

The desired outcome is (i), not (ii); that has always been assumed without discussion.

WELL, THIS IS A BIT HASTY AND AUTHORITATIVE. FOR INSTANCE, John Drury AND Colling Philips HAVE BEEN ARGUING FOR YEARS THAT SOMETHING VERY MUCH LIKE (ii) IS WHAT HAPPENS. PERHAPS BOTH SYSTEMS HAPPEN (INDEED, THAT WILL BE CHOMSKY'S CONCLUSION).

Thus if $H = T$, $XP = VP$, and β is either an expletive merged to TP or a DP raised from XP, the result of Merge would be (58i), with β the SPEC of T and VP remaining its complement, not (58ii), with β becoming the complement of T and VP its SPEC:

(58)

- (i) [T β [T VP]]
- (ii) [T VP [T β]]

RIGHT, NOBODY WOULD SAY THAT, OF COURSE. IF YOU PROCEEDED TOP DOWN YOU WOULD DO THINGS, FOR INSTANCE, AS DRURY DOES THEM, MERGING, FIRST, THE EXPLETIVE TO T, AND THEN ADDING VP. YOU HAVE TO BE COOPERATIVE IN THE GAME...

But the reasons are not entirely obvious.

THAT'S WHAT'S IMPORTANT, AND THE FIRST COMPLETELY FAIR POINT IN THE LAST FEW SENTENCES.

Each outcome satisfies a reasonable condition: (i) satisfies the familiar Extension Condition (55a); (ii) satisfies the condition of Local Merge (55b).

RIGHT. SO THE MINIMALIST CANNOT JUST GO WITH (i) BECAUSE OF TRADITION, OR EVEN BECAUSE 'IT WORKS'. ONE HOPES TO BE AS HONEST WITH THE CHOICES ELSEWHERE, EVEN WHEN IT ISN'T OBVIOUS THAT WE WILL IMMEDIATELY GAIN SOMETHING.

One possibility is to stipulate that the Extension Condition always holds: Operations preserve existing structure.

IN R&R THE DIALOGUE GOES IN THIS DIRECTION, TRIGGERED BY THE OTHER. THE CHARACTER CALLS IT THE 'GRANDMOTHER PROBLEM'. YOU GO BACK IN TIME AND KILL GRANNY; THEN WHAT? THE PROBLEM IS THAT WE'RE NOT IN THE PHYSICAL WORLD HERE (AT LEAST NOT OBVIOUSLY OR NOT DIRECTLY), AND IT IS COMPLETELY UNCLEAR WHY YOU SHOULD HAVE TO PRESERVE EXISTING STRUCTURE.

Weaker assumptions suffice to bar (57ii) but still allow Local Merge under other conditions. Suppose that operations do not tamper with the basic relations involving the label that projects: the relations provided by Merge and composition, the relevant ones here being sisterhood and c-command.

AGAIN, THIS SEEMS STIPULATIVE, ALTHOUGH PERHAPS IT RELATES TO SOMETHING PROFOUND ABOUT THETA THEORY. I SUSPECT THIS IS PROBABLY RIGHT, BUT DON'T REALLY KNOW HOW TO MAKE IT FOLLOW NATURALLY WITHIN PRESENT ASSUMPTIONS. WE SHOULD PERHAPS TRY TO RETURN TO THIS WHEN WE EXAMINE THETA RELATIONS IN MORE DETAIL, WITH AN EYE ON UNDERSTANDING WHY WE CAN'T 'TAMPER WITH LABELS'.

Footnote {New containment relations are defined whichever choice we make, in incommensurable ways;

I SUPPOSE THIS IS A JOKE, THE 'INCOMMENSURABLE' WAYS BIT.

identity is irrelevant.}

THIS, I DON'T UNDERSTAND. SOME OF THESE CRYPTIC COMMENTS ARE OFTEN WRITTEN RESPONSES TO QUESTIONS THAT PEOPLE RAISE, AND THUS UNDERSTOOD ONLY BY WHOEVER ASKED THE QUESTION. IF THAT PERSON IS READING THIS, PERHAPS THEY COULD GIVE US A HINT.

Derivations then observe the condition (59), a kind of economy condition, where R is a relevant basic relation:

(59) Given a choice of operations applying to a and projecting its label L , select one that preserves $R(L,G)$

THIS IS THE 'NO TAMPERING WITH LABELING' CONDITION, WHICH WE HOPE TO DEDUCE.

In the case of (56)-(57), the principle (59) selects (i) over (ii). In (i), basic relations of H are preserved, but not in (ii), which modifies sisterhood and c-command for H (in fact, the category $a = \{XP,H\}$ disappears altogether).

One case permitted under (59) but barred by strict adherence to the Extension Condition is head-adjunction. The standard assumption is that in this case Local Merge takes precedence over the Extension Condition. Thus in (56)-(57), (ii) is chosen over (i) for β a head H' . Adherence to the Extension Condition would change the status of H' to an XP; V-raising to T, for example, would create a VP-TP structure, with VP the SPEC of T, contrary to intention. The uniformity condition for chains (MP, 4.1 (17)) bars that choice, leaving as the only option head-adjunction yielding (ii), in violation of the Extension Condition. (59) is satisfied vacuously.

Footnote {Whether sisterhood (hence c-command) is preserved depends on how (or if) the notion is defined for head-adjunction. See note HADJ.}

Head-adjunction thus provides some reason to weaken the Extension Condition as proposed, permitting Local Merge if (59) is satisfied.

For XP-merger, (59) eliminates the ambiguity of choice of K for second Merge, imposing the Extension Condition (55a) rather than the Local Merge Condition (55b). But the choice remains open for third Merge.

THIS IS RICHARDS'S 'TUCKING IN'.

Suppose we have the outcome (i) of (56)-(57), as required, and now merge G to it (either by pure Merge, or with G extracted from XP). The two possible outcomes are (60i) (satisfying the Extension Condition) or (60ii) (satisfying Local Merge):

(60)(i) $\{G, \{\beta, \{XP, H\}\}\}$

(ii) $\{\beta, \{G, \{XP, H\}\}\}$

In (i) the newly merged phrase G is the outer SPEC, in (ii) it is the inner SPEC. Either way, sisterhood and c-command are preserved for H, satisfying condition (59). The discussion so far has kept to the Extension Condition (55a) (e.g., (8ii)). The question becomes more intricate when we cast a wider net, another issue deferred here.

Footnote {See Richards (1997) for evidence in support of Local Merge in the important case of multiple Move. See also note QURA; the proposal there falls under Local Merge, with the Extension Condition obviated vacuously for "post-cyclic" QR.}

Both Local Merge and Extension yield cyclicity and hence satisfy the optimality condition (52I), but in slightly different ways.

NOTE: CYCLICITY IS NOT EQUATED ANY MORE WITH THE EXTENSION CONDITION. CYCLICITY JUST MEANS 'WITHIN THE SAME CYCLE', NOT 'CREATE STRUCTURE IN THE OUTSIDE OF WHAT YOU HAVE'; THEN OF COURSE YOU HAVE TO CHARACTERIZE WHAT THE CYCLE IS, WHETHER PHASES OR SOMETHING ELSE (OR BOTH).

The distinction brings out an empirical difference between the derivational definition of c-command and the definition given earlier in terms of composition of elementary relations.

Footnote {See note DER; also Collins(1997), 84.}

Combined with Kayne's Linear Correspondence Axiom (Kayne 1994), as in the work cited, the derivational approach entails that the Extension Condition is inviolable.

I DON'T UNDERSTAND THIS. THE DERIVATIONAL APPROACH ENTAILS THAT, IF YOU VIOLATE THE EXTENSION CONDITION, YOU CAN'T BE LINEARIZED *BY THE LCA*. THAT MUCH, I UNDERSTAND. HOWEVER, WHY CAN'T IT BE THE CASE THAT WHEN YOU CANNOT BE LINEARIZED BY THE LCA ONE OF TWO THINGS HAPPEN: (I) YOU CRASH (CHOMSKY'S REASONING), OR (II) SOMETHING ELSE IS RESPONSIBLE FOR LINEARIZING YOU (E.G. HOFFMAN'S INFORMATION THEORY, OR INCOMMENSURABLY MANY OTHER WAYS). NOTHING IN THE LOGIC OF *WHY* THE SYSTEM USES LCA FORCES *THAT* PROCEDURE TO BE THE LINEARIZING ONE; IT'S THE MOST ECONOMIC, ARGUABLY, BUT IN GIVEN CIRCUMSTANCES THE MOST ECONOMIC MIGHT NOT YIELD CONVERGENT RESULTS. I KNOW, I KNOW: THIS IS CHAPTER 4 LOGIC. BUT STILL, IT'S NOT OBVIOUS TO ME WHY THIS ISN'T NOT ONLY MINIMALIST, BUT FURTHERMORE THE THEORY MAKING THE FEWEST ASSUMPTIONS.

The compositional approach still allows the option of Local Merge, as in head-movement and Merge in inner SPEC.

TRUE, BUT I THINK IT'S A PITY TO GET RID OF LCA, WITH ALL THAT IT ENTAILS (NOT JUST KAYNE'S RESULTS, ALSO, FOR INSTANCE, A PREDICTION ABOUT BINARY BRANCHING, AND OTHERS).

The steps towards minimalist goals discussed above improve the MP theory in other respects. It was argued there that Merge is preferred over Move, but on dubious grounds.

IN FACT, WORSE THAN DUBIOUS; WRONG. A TYPE OF OPERATION WAS CONFUSED WITH A TOKEN OF ITS APPLICATION.

Now the preference is immediate: Agree and Merge are each components of Move, so it is a simple matter of more vs. less.

THIS WAS ALREADY TRUE OF ATTRACT, THOUGH.

In fact, because of the extra component (II) of (54), Move is more complex than combined application of Agree and Merge. Thus if a derivation D has reached the stage (61) and an expletive is available, D must proceed to (62i) (Merge combined with Agree, the latter to establish the T-associate relation), not (62ii) (Move, then backtracking to the alternative (i) when the derivation crashes with an unused expletive):

FINE, ALTHOUGH I THOUGH THE 'BACKTRACKING' IN QUESTION WAS IN ANY CASE PROHIBITED (IN OTHER WORDS, SHOULDN'T IN THAT UNWANTED CASE THE DERIVATION JUST CRASH, PERIOD?)

(61) [TP T be [a proof discovered]]

(62)

(i) there was a proof discovered

(ii) a proof was discovered

Again, a look-ahead property disappears. This sharpens the analysis of (10i)/(12i): the desired outcome is not just an option, but rather the only option at the stage (61) of the cycle.

OKAY, BUT IF THE LOOK AHEAD OPTION ALREADY DISAPPEARS, WHY DO WE HAVE TO STIPULATE IMMEDIATE DECISIONS? IN FACT, IT'S WORSE: EXACTLY WHAT CONDITION DEMANDS IMMEDIATE DECISIONS? (I CAN SEE THAT A NUMBER OF FACTORS CONSPIRE TO GENERALLY YIELD THE DESIRED RESULT, BUT THAT'S DIFFERENT FROM STIPULATING IT, AND WOULD LEAVE THE DOOR OPEN TO INSTANCES WHERE, IN FACT, YOU GET A MINIMAL LOOK AHEAD WITHIN A PHASE.

Suppose we have reached the stage (61) and the only unused element of the lexical subarray is the complementizer C. Since Merge is preferred to Move, the theory of MP wrongly predicts merger of C and TP, barring dislocation of "a proof" to the subject position to yield (62ii)

(Eduardo Raposo, pc).

RAPOSO HAS TO TELL US EXACTLY HOW THAT WAS A PROBLEM FOR CHAPTER 4, SINCE I DON'T SEE WHY, AT THE TP LEVEL, THE MOVEMENT ISN'T REQUIRED BY CONVERGENCE DEMANDS (EPP FEATURE).

The problem is overcome when we recognize the asymmetry of Merge. T contains no selector that allows C to merge to TP. With uninterpretable features of T deleted in the strict cyclic derivation of full TP (see (53)), the derivation can move on to C, which selects TP, projecting C.

In MP it is speculated that categories lacking interpretable features should be disallowed; specifically AGR, consisting only of uninterpretable f-features (chap. 4.10). That conclusion is forced in this version.

Footnote{In MP, it could be avoided only by recourse to the (dubious) distinction between deletion and erasure. Nothing follows about functional categories with semantic properties, as in Rizzi (1995), Cinque (forthcoming).}

Suppose α is an LI that consists of uninterpretable features only and selects β , yielding the syntactic object $K = \{\alpha, \beta\}$ with label α . In the course of a convergent derivation, α will disappear, leaving K and higher projections of α without a label. But terms without labels are not well-formed syntactic objects.

KEEP THIS IN MIND FOR LATER ON, AND NOTICE THE TERMINOLOGICAL NIGHTMARE. HERE 'SYNTACTIC OBJECT' IS USED IN THE CHAPTER 4 SENSE. PREVIOUSLY WE WERE USING IT IN THE NEW, CYCLIC SENSE.

Accordingly such elements as AGR not only might not exist, but cannot exist, on rather plausible assumptions. The argument carries over to other cases, among them semantically-null determiners D_{null} . If true D relates to referentiality/specificity in some sense, then an indefinite nonspecific nominal phrase ("a lot of people," "someone" that enters into scopal interactions, etc.) must be a pure NP, not DP with D_{null} (and EPP cannot be stated as a D-feature). Notice that the argument holds only for heads α that select β . If α enters a derivation by Merge to a selector, deletion of α leaves a legitimate object (thus expletives might, in principle, delete completely).

WHAT ABOUT 'DO' AND THE AUXILIARIES THAT DELETED IN CHAPTER 3?

We have explored a number of lines of argument in the attempt to refine and improve the minimalist program, including (I) restriction of basic operations to Merge and Agree (the latter reducing to deletion of uninterpreted matched features), each satisfying conditions of "good design," each preempting the more complex operation Move; and (II) restriction of access to the features F provided by UG in successive steps ((3) and its extension to cyclic derivation based on lexical subarrays). Basic properties of CFCs are then accommodated. These and other cases discussed reduce operative complexity in a natural way, reinforcing the suspicion that there is something to the curious principles (28) and the intuitions and architectural conceptions on

which they are based. Basic relations are restricted to those provided by Merge and composition.

NOT OBVIOUS TO ME WHAT 'AND COMPOSITION' MEANS

A number of other ways of overcoming errors and defects of MP and advancing the project of MP and related work have been outlined. We have raised, but surely not solved, the problem of whether the most striking apparent imperfections of narrow syntax, violation of the Interpretability Condition and dislocation, are true imperfections or are reasonable ways of satisfying design conditions, perhaps with uninterpretable features serving as the mechanism to induce structural properties required by interpretive systems at the interface. The discussion has been largely theoretical, with only a few indications about how things turn out on "best possible" expectations. Many questions have been left dangling, innumerable others unmentioned. I hope to return to some of these topics in the continuation.