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,ß.

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 ARCANE 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 URIA.}

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 DINAMICS 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 noncategorial 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, PORTUGUSE 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 ENLISH 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 INDENTICAL 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 SIGURDSOON'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:

Sigurdsson 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 TRIVIALLY 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 TRIVIALLY. 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 a ß G, WHERE IS COMMAND.

the [Q] feature of the already checked wh-phrase (ß 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 [ß Hß...a], Hß a CFC and ß minimal,

(a) if Ha is C, Hß is independent of a

(b)if Ha is v, Hß = Tß agrees with EA, which may raise to SPEC-Tß though XP cannot

(c) if Ha is Tdef, if Hß is T then XP raises to SPEC-Tß if there is no closer candidate G for raising; and if Hß 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 Ta. We can limit attention to Ta 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 Ta; ASSOC is either raised to SPEC-Ta 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 Hß 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 Tß. These can only be deleted by agreement with EA, deleting its structural Case feature as well, with EA either raising to SPEC-Tß

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 Tß-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 [Tdef 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 Hß which is either v (ECM) or T (raising). Suppose Hß 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

EPP-feature (see note BAR).

TO REPEAT, THIS IS A SERIOUS DIFFERENCE WITH CHAPTER 4, WHERE RAISING TOOK PLACE UNIVERSALLY, EVEN IF IN SOME LANGUAGES IT WAS AT LF, VIA PROCRASTINATION.

If Hß is T, XP raises to SPEC-Tß unless barred by a closer candidate G. If Tß is nondefective, it either agrees with XP and checks its Case (nominative) or shifts to default, as already discussed. If Tß too is defective, then XP raised to SPEC-Tß will have to be associated with a still higher T or v, by raising or agreement in situ.

The basic properties (50) of CFCs therefore follow from simple and plausible assumptions. Without running through cases, it should be clear that the phase-impenetrability condition (21) holds for A-movement for the same reasons (see p. 00).

BRACE FOR THE FOLLOWING PARAGRAPH; IT TOOK ME ABOUT AN HOUR TO FIGURE OUT!

Further insight into these matters should derive from raising constructions with quirky Case moving to matrix subject, as in (51):

Footnote{Translations of Icelandic examples: (i), with a default variant, from Sigurdsson (1996); (ii)-(iii) from Boeckx (1997). See also (15), (45iic), note PERS.}

(51) (i) me(DAT) thought(pl) [t{me} [they(pl,NOM) be industrious]]

(ii) *me(DAT) seem(pl) [t{me} [John(DAT) to like horses(pl,NOM)]]

(iii) *John seems(sing) me(DAT) [t{John} to like horses]

LET ME START BY TRANSLATING THE EXAMPLES INTO A LEGIBLE NOTATION, USING DASHES TO INDICATE MOVEMENT, DOTS TO MARK AGREEMENT, DEFAGR TO MARK DEFAULT AGREEMENT AND KEEPING THE NOTATION 'DAT' TO SIGNAL QUIRKY CASE--AND REMEMBER, ONLY THE FIRST EXAMPLE IS GOOD:

```
                             ................
                            :                :
(i) [me-DAT thought-AGR [  t    [they-AGR be industrious]]]
   IP  |               vP |  CP
       |_ _ _ _ _ _ _ _ _ |
```

```
(ii) * [me-DAT seem-DEFAGR [t [John-DAT to like horses-AGR-
NOM]]]
      IP  |            vP  | IP
          |_ _ _ _ _ _ _ _ |
```

```
              .......
             :       :
(iii) * [John seem-AGR  [ me-DAT [  t to like horses]]]
       IP |             vP      IP   |
          |_ _ _ _ _ _ _ _ _ _ _ _ _ _|
```

(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.

WHICH DOESN'T MEAN, INCIDENTALLY, THAT THOSE FURTHER CONSTRAINTS WOULDN'T BE NATURAL. THE ISSUE IS:

Footnote{Unwanted chains could be formed without further constraints: e.g., given a and ß, and G in a, G could be merged to ß (by Move), which is then merged to a, forming an unwanted G, G chain violating conditions.

```
THAT IS:  ß,    α        VIA MOVE:  ß <--> Γ    α
             /     \                       | /     \
            /...Γ...\                      |/...t...\
                                              |_ _/
                 XP/αP
AND THEN:         /  \
                 X    \
               /    \  \
             ß       Γ   α
                     | /     \
                     |/...t...\
                        |_ _/
```

IT IS NOT COMPLETELY CLEAR TO ME THAT ALL POSSIBLE VALUES OF X YIELD BAD RESULTS. FOR INSTANCE, NUNES'S SIDEWARD MOVEMENT, OR BOBALJIK AND BROWN INTERARBOREAL MOVEMENT IS PRECISELY OF THIS SORT, WITH X=ß. ONE MIGHT WANT TO ELIMINATE THIS KIND OF MOVEMENT FOR PRINCIPLED REASONS, BUT IT SHOULDN'T BE BRUSHED OFF, PARTICULARLY GIVEN ITS EMPIRICAL PAY OFF (NUNES, HORNSTEIN).

Motivated by Pollock's split-inflection ideas, chap. 2 of MP adopts Attract-a in addition to Move-a. Chap. 4.5.6 seeks to eliminate Move-a as an independent operation, taking it to be conditional on Attract, revised as feature attraction. Considerations reviewed here seem to me to support the general point of view (now much simplified).}

THE SPIRIT OF ATTRACT REMAINS, CODED IN TERMS OF AGREE. HOWEVER, CERTAIN PROPERTIES OF MOVE HAVE NOW BEEN RENEWED (EPP EFFECTS), AND I WOULDN'T BE SURPRISED IF SOME PROPERTIES OF THE SORT PESETSKY HAS BEEN STUDYING (FOLLOWING WORK BY RICHARDS AND OTHERS) HAVE TO BE EXPRESSED IN TERMS OF MOVE, AND NOT ATTRACT.

SHIFT TO A DIFFERENT TOPIC ALTOGETHER, EMPHASIZING CYCLICITY:

A crucial property of deletion is that a deleted feature is invisible at LF and inaccessible to CHL (the [±active] property), but accessible to the phonological component. This property poses a problem on the assumption of MP that Spell-Out applies at a single point in a derivation: pre-Spell-Out, the probe must delete when checked but yet remain until Spell-Out.

Footnote{The problem is noted there but left unresolved: chap. 4, note 50.}

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, ß) 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,ß. Minimally, K should consist only of a and ß, so K = {a,ß}. 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 sparest 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 label(a) = a, for a 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,ß} and the ordered pair a,ß, a adjoined to ß.

I MUST STEP HERE AGAIN WITH ANOTHER FORMAL POINT, THE ORDERED PAIR a,ß IS SOMETHING LIKE THE SET {{a}, {a, ß}} (THE ANGLED BRACKETS ARE JUST A NOTATIONAL TRICK). I FIND THIS SOMEWHAT INTRIGUING, BECAUSE BY THE MERE FACT THAT YOU HAVE TO LINEARIZE {a,ß} INTO {{a}, {a, ß}} 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 ß AND THUS OBTAINING {{a}, {a, ß}} 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 WETHER 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,ß}} (substitution) or {G,a,ß} (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 ß. 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,ß by general principles, then the result of merger of a,ß is simply {a,ß} or a,ß.

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 ß to form a,ß. Given the asymmetry, it is natural to conclude that the adjoined element a leaves the category type unchanged: the target ß projects. Hence adjunction of a to ß forms K = {G,a,ß} where G is the label of ß. Eliminating redundancy, the operation forms K = a,ß.

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,ß 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,ß), as can be seen by reviewing cases.

Set-Merge of (a,ß) 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 WETHER 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,ß} or a,ß 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 ß, 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 ß as a candidate for "pied-piping"

(III) ß 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 ß 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) ß 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 ß to a. In this case, ß 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) {ß,{XP,H}} (as in (55a)) (ii) {XP,{ß,H}} (as in (55b))

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

```
(57)
      H                      H                            H
    /   \      --> (i)     /     \         --> (ii)     /    \
   H    XP                ß       H                    XP     H
                                /   \                       /   \
                               H     XP                    H     ß
```

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

(58)

(i) [T ß [T VP]]
(ii) [T VP [T ß]]

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), AD 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 ß 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,{ß,{XP,H}}}

(ii) {ß,{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 PITTY 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 a is an LI that consists of uninterpretable features only and selects ß, yielding the syntactic object K = {a,ß} with label a. In the course of a convergent derivation, a will disappear, leaving K and higher projections of a 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 Dnull. 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 Dnull (and EPP cannot be stated as a D-feature). Notice that the argument holds only for heads a that select ß. If a enters a derivation by Merge to a selector, deletion of a 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.