

At the GLOW talk, I discussed some conceptual problems that arise in a theory of the OB type. I suggested another approach which I suggested would overcome to some degree, sometimes completely, sometimes not, these conceptual problems. I also mentioned that this approach has different empirical consequences: It's incomparable in empirical coverage in that it properly explains some things which were not explained in the other theory, but it doesn't cover some of the material in the other theory. I then suggested that that was a good result since the things covered in this theory seem very central whereas the things covered in the other theory and not in this one seem rather peripheral. We may look to the future Markedness theory to justify these differences. I'll assume that this approach is correct and I'll pursue this alternative now, presenting it in a slightly more precise form. This alternative theory borrows quite wildly from a lot of current work. I think it ends up being extremely simple and quite elegant, that is, it has nice conceptual properties. That's what I want to look into. In doing it, as always, it's worth making a distinction between some sort of leading ideas and the execution of these ideas. As everything thing is hopelessly underdetermined by the evidence available to us, even available to people that know more languages than I do, it means that in executing some of these ideas it's necessary to make a fair number of arbitrary decisions, that is the leading ideas allow a range of possibilities of execution. For concreteness I want to give more or less precise instantiations of the leading ideas. But in the course of the discussion I'll try to distinguish between cases where it's a matter of execution to make a decision in order to proceed, where there are alternatives in the framework of the leading ideas, and these cases where the leading ideas are at stake.

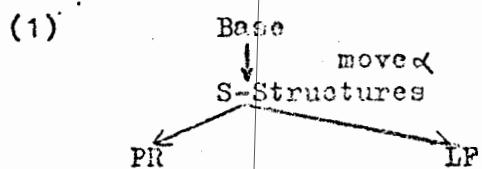
In the present phase of linguistic theory, one of the nice properties that is the range of phenomena that has been shown to be relevant data is so substantial and the intricacy of the theory has reached such a point that any slight change that you make has very wide, often unpredictable, consequences elsewhere. That's nice, because we know that that's what the theory ought to have. Of course it's bad too, it raises difficulties for research because when you start fiddling somewhere very odd things happen somewhere else. It may well be that the path that I try to choose among the various alternatives isn't consistent. If so, then I'll continue to change it as we proceed, as always.

I am assuming some kind of theory of UG, some sort of system with parameters. When we fix the parameters we derive one of the core grammars. An interesting property of the kind of theory I am thinking about is that, without exception, they only permit a finite number of core-grammars. That is if you look at UG and ask where are the possibilities of expanding it to an infinite set, well they really don't exist: It doesn't exist in the lexicon, obviously, assuming a rational theory of

morphology, except in the sense that words can get longer and longer, which is an interesting set. It doesn't exist in the component that relates to LF if it's more or less correct to assume that these principles are not subject to very much parametric variation. That makes a lot of sense because there's simply no evidence to fix them in that domain, except very marginally. There can be variations in the relation to Phonetic form, surely languages have different phonologies, but we're all assuming that phonology is kind of a finite system. If the relationship between underlying structure and surface structure is some very weak T-rule, like move~~α~~^{to speak of}, or some equivalent interpretive rule, then that ~~doesn't~~^{any} allow any variation. So we're left with the base, and assuming some variant of the X theories that have been proposed is correct they all only have a finite number of options. Hence it follows that there is only a finite number of core-grammars. That has a lot of bearing on a lot of quasi mathematical work going on on questions of learnability and generative power. In fact it trivialises those questions if it's true: There's no theory of learnability if there's only a finite number of core-grammars. That doesn't say that the work in learnability is pointless. It's saying that it's studying a very high idealisation: It must be kicking some element of UG and say "let's pretend it isn't there, so there could be an infinite number of these systems". Then let's see whether the system has some sort of property. That could be interesting, but one has to understand it in these terms. The same is true of problems of recursive generation. There's been a lot of literature in the last couple of years since Peters and Ritchie's badly misunderstood paper; It has been taken by lots of linguists to suggest that anything goes because languages aren't recursive and so on. Now the absurdity of this interpretation becomes transparent if in fact there's only a finite number of core-grammars. It might turn out ~~that~~^{the} languages that they generate are all non-recursive, in fact not even recursively enumerable, in fact arbitrarily wild. This would have no significant consequence at all with regard to questions of learnability or use of language or anything else as far as I can see. It's worth thinking about whether in fact the proper theory should permit only a finite number of core-grammars. It seems to me that the actual theories that people are working with do have that property, and it's probably the right property.

I am going to assume, to begin with, the framework of OB or F&C or other recent papers. This means that there's a base that generates base-structures of some sort. There's a transformational rule "move~~α~~" which generates a set of structures for which I'll make up a new name, I'll call them "S-Structures" suggesting "Surface-structure" but not using the term. Then there's some mapping onto Phonetic representation and some mapping onto Logical Form (LF), with the usual provisos, that is the properties of LF are to be determined empirically and will not be settled by some extrinsic concern such as formalising inference or something like that. So we have (1) next page. I suggest, as a minor terminological point that it would be a good idea that we avoid the term surface structure, or, if we use it, to use it in the old original sense. The only reason for this is that the use of the term "surface structure" has caused a fantastic amount of confusion, in particular among linguists.

What has happened in the last couple of years is that the term "Surface Structure" has been changed in extension. Whereas in early transformational work it referred to something close to PR, it now refers to something that would have been called deep structure earlier. The term "Surface Structure" may suggest that some radical change has taken place about the relation to surface structure and semantic representation. In fact no such change in any significant sense has taken place because everyone agrees that what used to be called surface structures are not the proper object for semantic representation. So if you don't know what the concept really means it's very easy to be confused. And there are papers by people like T. Langendoen or J. Searle that show much confusion of this type. So I think it might be useful to suggest a different term, although all terms are misleading, maybe this one will not be misleading for two years or weeks or something. For a clear presentation of this take a look at Koster's book where the issue is discussed. So we have something like (1):



Now let's have a look at the various parts. As far as the base is concerned I'll be assuming some version of the X-bar theory. The version of it I'll assume has the categories assumed by classical grammar, the basic category "substantive", including nouns and adjectives and the basic category "predicate" or something including verb and adjectives. So we have a system based on the features $\pm N$, $\pm V$ where $+ N - V$ is noun, $-N + V$ is verb, $+ N + V$ is adjective and $-N - V$ is preposition. I am also going to assume, contrary to Jackendoff, that the lowest rule for categories that have lexical content is something like

$$(2) \quad \bar{X} \longrightarrow \dots X \dots$$

where "..." is fixed through all categories, that is I am assuming that in the unmarked case V, N and A have the same complement structure. That means that (3) and (4):

(3) destroy the city

(4) destruction of the city

are the same form in the base. Similarly (5) and (6)

(5) write the book

(6) writer of the book

and at the level of adjectives (7) is derived from (8)

(7) proud of John

(8) proud John

that is I assume that Adjectives have close to the full range of verbal and nominal complement structures. Now of course that's not true at Surface structure the reason being case-theory. I'll be very heavily relying on Vergnaud's idea that case theory is central to this enterprise. Since I am assuming that V's and P's

delete. Of course this then bears on the question of the structure of COMP. As I said the reason for that structure was mainly the formulation of the free deletion rule in COMP. I'll stop at this point although one might pursue the consequences further. Notice that this has lots of consequences. So if you have structures like

$$(12) V \dots (\underset{\text{COMP}}{\alpha})(\dots \text{INFL} \dots)$$

and if you want to state the selectional restrictions between matrix verbs COMP's and INFL's there will be connections between INFL and COMP (tense and that). There will also be connections between V and COMP. But now there will also be connections between the matrix verb and INFL because COMP may well not be expanded. That's a consequence that doesn't bother me much because selectional rules pick heads of categories and if you think of INFL as the head of S then it's not surprising.

Now let me talk about the various subcomponents of this theory. I have talked about the base and the general structure of the \bar{X} theory. The next point is the theory of Government. This is going to be rather crucial in a theory that uses case as a central notion. We are going to need a general notion like α governs β . Let's say to start with

$$(13) \alpha \text{ governs } \beta \text{ iff } \alpha \text{ minimally c-commands } \beta$$

α minimally c-commands β =_{def} α c-commands β and there's no γ such that α c-commands γ and γ c-commands β and not γ c-commands α

In other words the cases that we want to include in this are for example that if we have V NP NP the V minimally c-commands both NP's. That's the standard cases that we want for government. Of course that's very much keyed to the assumption that we have a configurational language, i.e. a language of the NP VP type. I don't know what happens in the VSO languages, or how the theory of government works out for those. There's no point for me to make a proposal because it depends on very intricate questions concerning the whole complex of things having to do with the NIC that nobody looks at from the point of view of the survey-type universal grammar. So we're in the dark as to what happens in the VSO languages, and what happens in those languages, like Portuguese, where you get Inflectional structure in the infinitives. It may be that the system works in a radically different fashion for VSO languages. It's probably the case that any definition of government in terms of c-command is probably superfluous for these languages. Anyway I want to add to (13) a definition of what α can be. I want to say that α is a lexical category or tense. In other word I want to say that NP doesn't govern but N does. Similarly A, P, V govern, i.e. we have

$$(14) \alpha = \pm N \pm V$$

Now what about tense government? That's almost a notation: You could drop it by saying that nominative is a property of certain types of structures, namely tensed clauses. Again that's a matter of execution.

One thing we want to capture is that in a structure like (15)

that is the - N categories are the case-assigners it follows that anything that is in the "..." of an X ≠ - N it's not going to receive case in the normal fashion but by some other way, one of them, typical of English-like languages, that use prepositions as the equivalent^b case structures, is to put in an empty preposition which has no semantic content and in fact no content at all and is just a case-marker that's just introduced to save X-bar theory. That's completely different from the way Jackendoff looks at it, but I think it works and there's a rather nice paper by Mona Anderson in NELS 8 in which she points out rather persuasively that if you assume this and the standard "move^c" rule you'll get movement in NP's of the kind you get in the city's destruction and so on precisely when you don't get of-insertion. What lies behind this is the case filter, Rouveret and Vergnaud's suggestion, which says that if you have something which has a phonetic matrix and it doesn't have case then the structure is out, i.e.

(9) *_{NP} (phonetic matrix) if NP has no case.

From (9) it'll follow that you're going to have in a language some sort of device to save the full range of X-bar structures. In English it's the of-insertion rule and in Italian, although I hesitate to talk about Italian it seems to me that there are some arguments for a "di-insertion" rule, in structures that have very similar properties. Hans den Besten has pointed out to me that German has rules that assign genitive in a marginal class of adjectives. So that might be a device to save the A+NP structure. You might expect a language to have some sort of device to preserve something like the full expressive power of \bar{X} theory in the face of (9) which I take to be universal. Continuing to look at the base I am assuming something like

(10) a $\bar{S} \rightarrow \text{COMP } S$

b $S \rightarrow \text{NP INFL VP}$

for languages like English French and Italian, obviously there are different kinds of languages that have very different properties. INFL is the inflectional system which contains \pm tense (+ Tense being tense and infinitive and having a way of ending up on the verb) and AG (agreement) which contains the features person, number, gender and case and also very often ends up on the verb.

As far as possible, within the \bar{X} theory, I'll try to keep to the idea that there are certain contexts, which ought to fall out of \bar{X} theory, in which you are going to have clauses, i.e. \bar{S} . In those contexts you ought to get all types of \bar{S} , tensed sentences and infinitives. That's what's going to happen if you let the rules run freely. Furthermore there shouldn't be any way of getting part of these structures elsewhere. There shouldn't be to VP appearing in NP's or something like that, because that destroys the simplicity of the \bar{X} system. The simplest thing is that there are positions where there are clauses and there are two types of clauses. Period. In some cases I think there's good evidence that that's true. So take indirect questions where you have an overt infinitive only but you get wh-movement which tells you it must be an S because surely you don't get wh-movement in VP's.^{Why not?} There will of course be lexical constraints that will vary from language to language as to whether, say, the

word for order will take an infinitive or a tensed clause or both or something else. That you can't get around: Languages differ at the level of lexical constraints, but I am assuming that that's some kind of independent grid imposed on it from the lexicon.

What about the structure of COMP now? Here I will be assuming a theory like OB an earlier papers that assumes that COMP has two positions, one that's filled with wh-phrases, one that's filled by \pm WH where - WH = that and + WH is the thing that goes along wh-phrases where they are interpreted as question. + WH may in fact be base-generated in words like whether. Also, in English, and that's very marked, for. I don't know of any other languages that have that, except modern English. So we have

- (11) (COMP wh-phrase \pm WH)
for

the order is not very important here: In languages that allow doubly filled COMP's you can find examples of both orders. Also there may be more structure, for example wh-phrases might be adjoined to S or \bar{S} . I'll keep to (11) as a notation for expressing the facts. This is a matter of execution.

(Question by someone as to the place of other conjunctions in the system in (11) like while. Chomsky says that true conjunctions or even things like prepositions like than and as don't enter the system in (11). If is probably in (11) but it's not a true conjunction.)

In the F&C theory there was very good reason to assume that the wh-phrase was really in the COMP. The reason for that was the rule of free deletion in COMP, which had nice properties and was intimately connected with the *NP to VP filter. In the OB theory there was a proposal that wh-phrase is deleted as much as possible up to recoverability (similar to an analysis for French by Ritchie). And I claimed there that if you did that you could get rid of the NP to VP filter. Notice that that has a consequence with respect to free deletion in COMP; namely the motivation for it is gone: There's no more motivation for the free deletion in COMP once you have a general principle that applies to wh-phrases anyway. Then we may go back to an other idea that was no good on the F&C system but is perfectly fine in the OB system, namely that the rule that sometimes expands COMP sometimes just doesn't apply, i.e. in the base we sometimes don't get that or for. That has the effect of the free deletion of that and for. Of course if you make that change it has all sorts of consequences all over the place, but I'll assume it anyway, that is, in tensed clauses, or infinitives you might take for, or that, none, and so on. That eliminates any need for the deletion in COMP rule, except for the question of what happens to the traces in COMP. I'll be assuming that there's a convention that says that uncase-marked traces just delete. So what we have finally is three principles, each of them fairly plausible, deletion up to recoverability, optional expansion for base rules and that principle that's really a special case of deletion up to recoverability, since it amounts to treating case as a real feature: If you have a phonetic matrix you can't

(15) (_{VP} V)

governance should be closely related to the strict subcategorization of the verb. A verb is going to strictly subcategorize in various ways, an object, an indirect object, an S and so on and it ought to govern the various things that it subcategorizes and nothing else. In fact we might even say, I'll come back to that later on, that it's the strict subcategorization features that do the governing, and that would carry over to N's and A's in the \bar{X} theory I am assuming since they are going to have the same \bar{X} structure they'll govern exactly as verbs do, though they won't assign case. Hence we could say that a category inherits its governance from its strict subcategorization features. Anyway the verb governs the things for which it is strict-subcategorized. So it follows that in a structure like

(16) (_{VP} V ... (_{PP} P NP))

where PP is part of the thematic structure of the verb, in Rouveret and Vergnaud's sense, that is strict-subcategorized by the verb, then PP will be governed but a NP within it will not be governed by the V. We could appeal to a locality principle in Koster's sense to explain why that should be. I will also assume that in (13) we add the proviso that there's no S' or NP between α and β , i.e. that S' and NP are absolute boundaries for governance. This has various consequences. So for example it leads us to believe that S' deletion is the device for representing exceptional case-marking. Again, this is execution but something has to distinguish believe and try and believe in English from croire in French. From this approach it follows that what will differentiate these cases will be S' deletion. It will then follow that exceptional case-marking will always be across S. Similarly there won't be any kind of case-marking across NP's. So in a structure like

(17) NP V (_{NP} NP's \bar{X})

there will never be any case marking between V and NP. Again one might appeal to a locality principle to explain that because the NP is already governed by Genitive but in this case I don't think it's going to work in this case because I'll suggest later on that that's optional and if it's optional there ought to be case-marking between V and NP. That suggests that making NP an absolute boundary for governors is the right approach. Other things follow from that because case-assignment is a special case of governance: You get case assignment when the governing category happens to be a case assigner.

(Richie notes at this point that there's still an asymmetry between S' and NP since there's an S' deletion rule but not NP deletion rule. Chomsky agrees and says that there are general principles that will tell you when S' is deletable and these principles have to do with I-containment, in a sense that I can't make precise. There are a number of things that converge toward such a conclusion: The adjectives that never take for infinitivals are the ones where you delete S', like certain, possible you don't delete the S'. So if we think that being able to

take for falls under the proper notion of L-containing. There's an other fact in English, originally noticed by Janet Fodor, namely that verbs that take an undeletable that are not bridge verb in Nomi Erteschick's sense. If we take S' to be a bounding category for subadjacency, always, and if we assume that the bridge verb phenomena is also related to S' deletion, then the facts would follow, if we assume that the undeletable that makes the bigger S' L-contain the smaller S. This does suggest something about NP, although I don't know how to work it out. For example one might argue that in some vague way indefinite NP's, the ones that allow the picture noun phenomena, kind of lose their NP boundary. This is really hand-waving. Hopefully some general principle will tell us why there's the asymmetry Richie mentioned)

What about headless relatives as in the Groos and van Riemsdijk theory? If I understand correctly we have structures like (18):

$$(18) V(NP(NP \alpha)(S, (COMP wh\text{-phrase}) (S \dots)))$$

What they show persuasively is that the wh-phrase is accessible, both to case-assignment and to strict subcategorisation. Can there be more than one governor for a unique NP for instance? As far as I can see the only case would be a structure like (19):

$$(19) V(S NP Tense)$$

if you do believe that bridge verbs involve S' deletion: In that case both the V and the Tense govern the NP. Plainly we want the Tense to be what counts, the NP is not going to get objective. Here again a locality principle might solve the problem. Assuming all that, we can now define a crucial notion namely

$$(20) \alpha \text{ is a governing category for } \beta \stackrel{\text{def}}{=} \text{there's some } \gamma \text{ such that } \gamma \text{ governs } \beta \text{ and } \alpha \text{ contains } \gamma.$$

So a governing category for something is just some category that contains its governor. So governing categories are kind of nested. If there's no multiple governors, governing categories are going to be nested sets, and those are going to be the relevant c-command sets for the binding theory. Now we may want a notion of minimal governing category, in the obvious sense, which we can define as

$$(21) \alpha \text{ is a minimal governing category} \stackrel{\text{def}}{=} \alpha \text{ is a governing category which properly contains no governing category.}$$

and we define α as being NP and either S or S'. I am not sure which, the difference being extremely slight, and only relates to for-infinitivals. For execution again I'll take S. So take a structure like

$$(22) (S \dots (S NP (VP V NP)))$$

in (22) both S's are governing categories for object NP, but the lower S is its minimal governing category. Now let me just state a fact: If NP is case-marked then the minimal governing category is the category in which it is case-marked. That takes care of the theory of government.

The theory of case comes next. Where are we going to assign case? again there are various options but at least by S-structure, and let's say, for simplicity's sake, at S-structure, by rules of the type shown in (23):

- (23)a. NP → nominative if governed by Tense
- b. NP → objective if governed by - N
- c. NP → inherently case-marked elsewhere (as determined by idiosyncratic properties of - N.)

Recall that (23)a could be restated as a structural property. As far as (23)b is concerned I am assuming that where objective is the unmarked case for non-nominatives. (23)c covers cases where a verb happens to take a genitive object. I am assuming that dativo is an inherent case-marking in this sense. In addition there will be special rules like the rule of of-insertion in English, or the rule for the genitive adjective complement in German, i.e. rules that are stuck in to let the X system work. I am going to assume also Richie Kayne's suggestion that in English the case-system has been lost, that is the subject NP will be nominative in tensed S's and the rest will be objective. There may be some marginal case assignment too. Now we also need some rule for the genitive case assignment in NP's so it might be something like

- (24) N → genitive in (NP N X)

I'll come back to that later on for more details. I also have to account that in cases of wh-movement we might have a structure like

- (25) NP (S, (COMP wh-phrase)(S ... (S, (COMP t)(S t))))

and the wh-phrase must inherit its case. This is a very marginal case and let's assume that we have to have whom if you have objective and you can't get whom if you have nominative. We want to say that it inherits its case from the variable it ultimately controls. So let's just say that, namely the c-commanding argument that it controls assigns case to the thing it dominates. We might wonder whether we want to assign case to the intermediate trace. There's no point because it has no phonetic effects. Also it has no effects on the case-filter. So we might assume that it's only the actual lexical thing that gets the case. So the condition might be that a lexical gets its case from its trace.

(Question by Jean-Roger as to how one can exclude *Who does it seem to have left and also exclude for instance the man that it seems to have left given the idea that (11) is a filter. Chomsky says that he'll come back to that later on)

Following Rouveret and Vergnaud's proposal I'll be assuming that participles don't assign case. In other words if passive participles are like

- (26) [-V] en (V kill) NP

then the NP can't get case because it's not governed by a [+V] element. That means that in a language that has that property, passive movement is obligatory, since the NP if it didn't move would ^{not} get case. Of course we could still have inherent

-10-

case assigned like in (27):

$$(27) (-v]^{en} (v)) NP_1 NP_2$$

NP₂ could get structural case. It wouldn't get it from the participle. So let's say that the inherent cases which are lexically determined may still be assigned and the positions to which they are assigned will be immune from passivisation, the reason being that the position has case, and it's going to turn out later that case-marked traces are always variables. Here they can't be variables because there's no operator, then passivisation can't apply to them. But the structurally case-marked positions won't be case-marked.

(Question by Dick Carter concerning languages where passive is marked by some special morpheme. Chomsky notes, in connection with Keenan's talk at GLOW that it's wrong to think of "Passive" as being a unitary phenomenon. You can define passive as ^{Ed. Keenan's} ~~that~~ structure in which when you put in individuals for the NP's you'll get synonymous sentences. Now the question is whether that defines a linguistic structure or significance. I don't think it does. Languages have all sorts of ways of having that phenomenon. There are two relatively pure cases; one of them is the movement case, English being an extreme example of it, where you have a pure structural rule: Anything can go into subject a subject position if no condition is violated. It doesn't have to have any grammatical relation with the verb, it doesn't have to be a referential expression, (you can have there). That's one mechanism that will give you something that meets Keenan's definition of passive. At the other extreme you can get pure morphological passives. There's a nice paper by John Mac Carthy on classical Arabic that discusses what may be an extreme case of this. The verb can be either have a passive or an active morphology. Depending on which you take you get a different series of cases assigned to the NP's. It turns out that the paired-structures meet Keenan's definition and also many other properties. I don't see any point in trying to get a notion of passive that includes this and the English case. Apart from Keenan's definition that they both meet they'll have radically different properties. In the pure morphological passive the pairing will always be narrowly constrained by the verb. In English, as has often been noted you get both cases. So you get the movement passive in things like (28):

$$(28) \text{John was taught French}$$

and the morphological passive in the so-called unpassive like

$$(29) \text{John was untought}$$

(28) and (29) have very different properties, in exactly the way that would be expected from this point of view: the "passive" "verb" needn't ever be a verb in (29), so you get things like John was unlettered. Of course (29) doesn't mean that nobody taught John. Similarly you'll never get French was untought John etc, just what you'd expect. As can be expected you get cases of ambiguity, as in

$$(30) \text{The door was closed}$$

Notice that you wouldn't expect the converse. You would never expect a language that has morphological passive to allow idiom chunks, free passive disregarding grammatical relations and so on. Now any approach to passive that will attempt to make the phenomena described unitary in terms of lexical insertion rules will miss all this, except by stipulation.

Question by Hans den Besten concerning the case assigned to the NP French in (28): Chomsky's answer is that one possibility is that the structural case is assigned to lexical items. Of course many other possible analysis are possible. Richie suggests that one might say that Past Participles assign case optionally. Chomsky agrees and notes that the rules as stated are ambiguous: Depending on which item you pick in (27) (i.e. [-V] or V) you might get case or not: If you do it optionally then it means that if you had a trace you're in trouble so you can't assign it in this case. If you have a NP left and don't assign it you're in trouble. Richie notes that if you assign inherent case in the base then you can get the PP stranding facts in French and non strandable languages for wh-mvt: If case is inserted in the base, then the trace will never be marked for case and it won't be a variable. Chomsky agrees, but notes that all is required for his present purposes is that case be assigned by S-structure.)

Now let me come back to Jean-Roger's question, namely the reason for the * in (31):

(31) *who does it seem (t to have seen Mary)

Well it ought to come out somehow that variables are going to be case-marked. Now there are well known problems with the OB theory, since if you have a relative clause structure like for example (32)

(32)*The man who John tried (t to see Bill)

deletion of who should render (32) OK because the case filter is not violated in (32). The obvious answer, which didn't make sense in the OB theory but which makes sense from this point of view, is to assume that the case filter is on the right. The reason is that in this theory government is the fundamental notion even at the level of logical form, so it makes perfect sense for the properties of case to be related to logical Form. Notice that it doesn't really matter whether you make the case filter apply at LF or at S-structure. If we adopt either of these, the problem doesn't arise because the wh-phrase won't have been deleted. Now if we state the filter the way we did at S-structure then it'll turn out that variables have case. What follows from this approach, as pointed out by Freidin and Lasnik is that EQUI is not going to be SELF deletion. I am not sure this is a bad result however because it makes Equi in English look more like Equi in other languages. This kind of situation as I said is typical: Anytime you change something somewhere, then lots of changes take place elsewhere. As for the difference that might arise from a decision to make the case filter apply at LF or S-structure one thing it relates to is May's QR rule:

the question is how are you going to interpret the quantifier part of the QR-rule: If that's like "for all x" then the problem of assigning it case arises. The point that we want to capture is that the variable position under QR gets case. You can think of various ways of doing it but the easiest way of doing it is to say that the case filter applies at S-structure. Then the question doesn't arise. If you did it at LF you would have a lot of fiddling around to make sure that the variable always gets case. Anyway it seems to me that the natural conclusion, subject to empirical disconfirmation, is that the case-filter applies at S-structure and is really just a property of case assignment.

(Question by Richie: How does it work exactly? Chomsky: Well it just says at S-structure you finish up the case assignment and then you ask yourself Have I done it properly? If I have violated the case filter then *.)

Now I want to point out that all this relates to what Wasow called the strong cross-over cases, like in (33)

(33) Who did he say Mary kissed
or the example given by Freidin and Lasnik

(34) Who did he say had won ~~that~~ X (^{he said} ~~X~~ won)

The strong cross-over cases show that the wh-variable behaves exactly like a name (variables left by QR also behave like a name) with respect to anaphora. That is they don't get antecedents, which is perfectly sensible if you think about it: It's extremely unclear what it would mean for a bound variable bound by its quantifier to be also bound by an argument. In fact it's natural to think of a bound variable as a name, that will be instantiated with names. The strong cross-over cases show that that's true with regard to the SSC and the NIC: That is being like names rather than anaphors they simply are not subject to either of these conditions. I will consider this as rather crucial evidence. So, in principle bound variables are going to be like names and are not going to obey any of those conditions.

(Question by J. Guéron concerning sentences like his mother loves John. Chomsky: That's a case of the weak cross-over, like who does his mother love. They are more complicated but they still show the same phenomenon, that is that bound variables behave like names. What's crucial in the first example is of course whether you get his mother loves JOHN or his mother LOVES John. I am assuming that in the first case you have a bound variable.)

Now of course we nevertheless have the fact that something like the NIC appears to hold for wh-movement. The relevant examples are of course of the type:

(35) *Who do you wonder (g, (how well)(t saw Bill))

which are hopelessly impossible even for people who accept similar cases with extraction of a non subject. So we seem to be in trouble since (35) seems to

show that wh-traces do behave like anaphors and yet (33) and (34) show that they don't rather than they behave like names. My solution to the problem, which I'll come back to (see lecture 3) is that the idea that (35) is a NIC violation is a spurious generalisation. So I'll be assuming that what was called the NIC in OB is not a unitary phenomenon. What I'll be calling the NIC from now on is the one that doesn't hold in (35), but rather the one that's involved in each other pronouns and disjoint reference. The thing that holds in (35) I'll just call the Residue of the NIC.

Notice that the case filter says intuitively that if a variable is case-marked for (NP e). We might ask whether that's an "if and only if". It may be that it is an "iff". But it would be plainly much nicer to derive that fact from other considerations even though it is true that variables are case-marked empty elements. By variable I mean an empty element which is bound by an appropriate operator. Note, as a point of terminology, that the word "bound" that appears when we talk about quantifiers and variables has nothing to do with the word "bound" that appears in the theory of Binding. Let's say that the notion "bound" in the theory of Binding is "Argument-Binding". It relates anaphors and their antecedents and so on. So something may be bound by a quantifier and not be "Argument-Bound" and conversely. Anyway going back to my definition for variable, the "appropriate operator" could be a wh-phrase in COMP, a quantifier in May's QR rule, it could be the "for x = John" operator that Richie talked about the other day. It could be the definite description operator. (Remark: by Richie that the fact that there are these two types of binding might be taken to be two realisations of one more abstract notion. Chomsky: Maybe but I think they are conceptually different. One is a logical notion, the other, binding by an antecedent, is a syntactic notion that enters the semantic interpretation. So I wouldn't take it as an advance if they were brought together.)

OK the next topic is the theory of indexing. One thing this theory will have to say is that we get coindexing with movement. I'll assume that. Now beyond that it's obvious what the simplest possible theory would say: Index freely. Let's therefore assume that that's correct and see if we get into too many problems. Of course we do get into lots of troubles, as usual. The thing to try to show is that the null hypothesis is nonetheless correct and that the reason for why you got into trouble was because you violated some independently motivated condition. Now it's clear that the general conditions we're all thinking about are going to rule out lots of the cases where you get into trouble. What you want to show then is that the other cases are also ruled out by something. There are cases that are not blocked by those conditions. So take a look at (36) which ought to be a sentence of Italian:

(36) l'uomo _S (wh-phrase)(mi domando _S chi(t vede i)))
(the man who I wonder who saw)

Now (36) is ambiguous and the trace can be a subject or an object trace. So if we index freely there's nothing that will block the indexing in (37):

(37) $l'uomo_{\exists} (\text{wh-}che_{\exists} mi \text{ domando}_{\exists} (S_{\exists} \text{ chi}_{\exists} S_{\exists} t_{\exists} vede t_{\exists}))$

which we obviously want to exclude. But there's an obvious condition that will block (37), namely the convention for inheriting case: the subject and object traces have different cases, nominative and objective respectively. If we state that case conflict is prohibited, whatever precise formulation we give, then (37) will be out. What I have just said will always work, except in a situation where a wh-phrase controls two traces that have the same case. So that's the only kind of problem that could arise, and if we index freely we might get the same index on the two traces. In cases where the case system doesn't uniquely pick out a case for every wh-movable position then independent principles might solve the problem, eg a) might say that a trace cannot both be argument-bound by a trace and bound by a quantifier (Jim Higginbotham's suggestion) or b) you could generalize case-conflict and say that two cases assigned to a wh-word, even if they are the same case generate a case conflict (Richie's suggestion) that is two different elements can't assign even the same case (which is using a notion like "government conflict"). (This is a very short summary of a very long and rather confused discussion J-Y Pollock))

Let's assume that the free indexing approach is right then we would want to avoid Jim Higginbotham's reindexing rule, that is things would be indexed once and for all. That raises a few questions to which I'll return but I think they are all soluble. The reindexing rule says approximately that a pronoun reindexes to a variable to its left except in some configuration. Higginbotham shows that this applies either after or before QR. Assuming that all he said is correct then we could restate it by saying that a derivation is all right if there is some point in it at which the pronoun is to the right of the variable. So we can do away with the need for a reindexing rule by adopting Jim's theory and thinking of it as a property of proper derivations.

Another thing that ought to be noted is that we are getting away from the spirit of Lasnik's theory. Take any case in which the noun and the pronoun are supposed to be completely unrelated by any anaphora connection. In Lasnik's theory it would mean that the pronoun is free in reference which means that it could refer to the NP or to some other NP. In the theory I have just stated it's going to turn out that they may or may not be coindexed, depending on how you did your free indexing at the beginning. In other words the pronoun could be obviative (not coindexed with the NP) or proximate (co-indexed). So I am not now using any variant of Lasnik's proposal. I am just using free indexing. We will want to say, just as a principle of interpretation, that if a NP and a pronoun are coindexed we interpret the pronoun as a variable bound by the NP. If two NP's are coindexed we will just interpret them as

coreferential or overlapping in reference. You'll get that with epithets.

Now let me talk about the theory of NP's. For the purpose of the present discussion three types have to be distinguished. Anaphors which I will take to be something that's lexically identified as an anaphor (each other for instance). So I assume that there's something in the lexicon that says perhaps because of its meaning, that each other has got to have an antecedent. So each other, reflexives, maybe some of the bound idioms in English (lost his way) are lexically identified as anaphors. Secondly I'll want things that I'll call ponominals. That's a descriptive term which will include pronoun and PRO. What's a pronoun? A pronoun is something that has a phonetic matrix and a bunch of features, person, gender and number. It may have case that it receives from somewhere else. Now PRO, contrary to what I was assuming for the last couple of years, is just a pronoun without phonetic matrix. One way of implementing this is to say that when you have structures like

(38). (NP (N—))

you can lexically insert the features of the pronoun and insert the phonetic matrix that's associated with the features. If you do, you get pronoun but if you don't you get PRO. So I'll be assuming that you don't have empty categories in languages except in very restricted conditions. So there will be some Empty Category Principle which will subsume the Residue of the NIC. PRO will not fall under this principle: They are not empty because they have features. So I'll be assuming that the relations of control are very much like the proximate interpretation of pronouns: That is the NP and the pronoun have to match in features just the way the NP and PRO have to match in features.

OK. Let's look at a problem that Richie pointed out to me the other day. Take a case of passive, for instance (39):

(39) John was killed

I am assuming a movement theory for passives for all the reasons I talked about earlier on. But now the question is: Why shouldn't we able to generate this in the base and then have an interpretive theory? That is why shouldn't there be an ambiguity in the interpretation. Well we don't want to have empty categories floating all around the place. So let's imagine that empty category never get indexed and that everything has to be indexed or else it's not well-formed. So part of the indexing theory says that you can only index things that are not empty. That means that only if you've done things by movement will you be able to index an empty category. Now if the representation of (39) when base-generated is (40): (40) John was killed (NP e)

then (40) is out. Now if the empty category is replaced by PRO then it's also out for reasons that I'll mention, namely that PRO will always have to be ungoverned by the Binding theory. If you have PRO in (40) it would be governed by the participle. So either way you can't have the interpretive option, which is of

~~something you want~~ value back to anaphoria. Let me say that I'll

continue to use the term obviative and proximate for all pronominals, for both PRO and pronoun, in the obvious sense. PRO is proximate if it's controlled and obviative if it's not controlled, that's what I was calling arb in OB. The theory of control is very largely devoted to determining where you can have obviative and non obviative interpretation of PRO. Notice that there was a serious problem in the OB theory. I said there that a PRO could get some index α that made it arbitrary. The trouble with that method is that it's in conflict with the fact that PRO's can be moved: When they move they're going to be coindexed with traces and then you're in trouble. So that theory didn't work anyway. That's just as well because we're going to get the same results with no theory in fact, namely free indexing will leave some PRO's uncontrolled, and they'll be obviative if they don't violate whatever conditions hold for obviative interpretation. That's all we need. So we have a much simpler theory. Notice in particular that we are not using anaphoric indices at all.

Let me talk now about the theory of Binding. So the theory of Binding has three principles A, B and C. The first principle is formulated thus:

- A. If α is an anaphor (i.e something lexically specified as needing an antecedent) or has no phonetic matrix (i.e trace and PRO) then (1) it is a variable
 - or (2) it is bound in every governing category (in particular in the minimal governing category if there's one)

A.(2) means in particular that each other has to be bound in all of its governing categories, in particular the smallest one. A seems to me to be the simplest statement of the facts. The second principle B can be formulated as

- B. If NP is case-marked, then

- (1) It is an anaphor (like each other that's going to be case-marked)

- or (2) It has to be free in every governing category

Remember the things that are case-marked are things with phonetic matrix or variables. Of these, first there are anaphors which have to be bound everywhere. What is left is vbls and things with phonetic matrix and those have to be free, that's all, they cannot be argument-bound (this doesn't mean quantifier-binding). Let's talk about pronominals now:

- C. If α is pronominal (i.e pronoun or PRO)

- then It is free in every minimal governing category

Intuitively speaking, C. means that PRO is always controlled from outside of its own S and it means that pronouns have disjoint reference.

Now what do I mean by bound? I'll define that as in (41):

- (41) α is bound (argument-bound) iff α is an argument coindexed with a c-commanding argument.

To say that α is free just means that α is not bound. Notice again that I am talking about argument-bound. So for example α might be argument-free but

nevertheless bound by a clitic. An argument is something that fills one of the main NP positions in an S, subject, object and indirect object.

Now go back to the case-filter. Recall what it said: Things with lexical content and variables have to be case-marked. Let's take a look at PRO. Let's imagine first that PRO is governed. Then there is a category, let's call it α , in which PRO is governed. Let's look at α and let's say in fact that α is the minimal governing category: If PRO is governed it'll have a governing category and obviously one of them will be a minimal governing category. So by case C, PRO has to be free in α . But by case A., since PRO is not a variable, (it has a feature structure) it has to be bound in every governing category in particular α . Now obviously this is a contradiction from which we conclude that PRO wasn't governed. Now there are various positions in which an NP is not governed, for example the subject of an infinitive, except under exceptional case marking. So in (42):

(42) It is unclear (_S, (what)) (_S PRO to do)
there's no governor (tense or lexical category) that minimally c-commands PRO. So that's one example of an ungoverned position and PRO will have to be in one of these.

Now let's look at trace. Suppose it is case-marked. Then it too has a governing category, let's call it α , in which it is case-marked. Now, looking at B, since it's case-marked, then it's either an anaphor, which it isn't, or free in every governing category, in particular in α . By A, it's either a variable or bound in every governing category. Now it can't be bound in every governing category (by B), therefore it is a variable. So we have the rest of the extended case-filter in fact. Notice that this tells you that if the trace happened to be the trace of NP movement, then you're in trouble because it's not a variable, since it's not bound by an operator. So it follows also that the trace of an NP movement cannot be case-marked, because if it is it has to be a variable which it can't be.

So that gives us the basic properties that we want, namely that the position for PRO is ungoverned positions. Notice that pronoun will never occur in ungoverned positions. The reason is that pronoun has a phonetic matrix, therefore it has to be case-marked, and if it's case-marked it's always governed. So there's something like complementary distribution between PRO and pronoun, although that fails in some cases, namely those situations where for one or another reason government is optional, then both PRO and pronoun might appear: You'll get structures in which you get pronoun if you pick the option of assigning case. You get PRO if you don't. But apart from those cases to which I'll return it's basically true that you have complementary distribution between PRO and pronoun.

So we have PRO is ungoverned, trace is uncase-marked, it may or

not be governed, but if it is case-marked it is a variable. It will also turn out, although I haven't said why yet, that trace must be governed. One odd last problem has to do with free relatives. Again assuming the Groos and van Riemsdijk theory. The structure we have is something like:

(43) (NP (NP e) t)

The outer NP is case-marked. So we happen to have a NP with nothing inside it that's case-marked. That would be a problem except for the fact that it's not true that it's an (NP e), it has NP's inside.

(Richie notes that a case not unlike that is the case of subject clitic adjunction in French.)

Chomsky, Pisa, Lecture 2.

(Question by someone having to do with notion "governing category" (see lecture 1 where following definition " α is a governing category for β " =_{def} there exists some γ where γ governs β and where $\alpha = \dots \gamma \dots$ and where $\gamma = \text{NP}, S$ " is given) followed by short dicussion concerning choice of S' or S in above definition. Chomsky notes that the only relevant strutures to make a choice are of the type:

(S they'd prefer very much (S , for (S each other to win))

Notes also that given theory of Binding, i.e

A. If α is an anaphor or lacks a phonetic matrix then

(1) α is a vbl or

(2) bound in every governing category

B. If NP is case-marked then

(1) an anaphor or

(2) Free in every governing category

C. If α is pronominal, then free in every minimal governing category (where "minimal governing category" is by definition an α which contains no governing category") then "you can unify the conditions for example for anaphors it's going to turn out that they have to be bound in every governing category which gives you both the specified subject and the NIC because of the way governing categories fall out, because of the very natural definition for governing category which says that the governing category for the case-marked elements is the category in which they get case, for the non case marked elements it's the ones that are structurally identical to those except that they don't have to be case-marked.)

Well let me say briefly what I was talking about last time. I gave a theory of the base hoping to be able to keep to what I think is the optimal theory that is the one that maximizes X-bar symmetries including in particular symmetries between the various types of clauses so we can keep to an optimal theory of the base. Then I talked about the theory of Government. The basic idea is that lexical categories (i.e V, N, P and A) govern the things that they subcategorize and other comparable things that they comparably c-command. Tense I am assuming governs what it c-commands, i.e the subject, although, as I said you can restate that as a property of tensed sentences. The theory of case says that for certain governing categories, namely verb and preposition, they happen to assign case. No later than surface structure (at surface struture, except for the case that Richie mentioned) in the usual way i.e Tense assigning nominative, the unmarked alternative is objective so that's what happens in verbs and prepositions and then there's the inherent case and Richie gave a good argument that it should be in the base. That would be prepositions in languages that have oblique

and verbs when they assign special cases. And then there are these odd things that a language may have to make it possible for the X-bar theory to be fully utilised, for example the of-insertion in English, and maybe these marginal adjectival things in German. If a language has such rules then the X-bar theory can be relatively fully utilised, that is, even the non case-markers will have the complements that the X-bar theory says that they should have. That's the natural way of thinking of rules like of-insertion.

The theory of indexing suggested in the first lecture gives up the whole theory of indexing of On Binding and basically reduces to co-indexing under movement and arbitrary indexing. That's a minimum. That means getting rid of anaphoric indices, of the notion "free_i" and the corresponding notion "designated index". If you can give all that up it's obviously a big improvement, because it does away with a lot of apparatus and deduces it from other principles.

The next thing is the NP theory. The relevant categories of NP's for this theory are lexically marked anaphors like each other or reflexives on the one hand and "pronominals" on the other. Remember that I want to modify slightly the notion "PRO". I want to say that a pronoun is a phonetic matrix plus a "PRO", where PRO is the various features that make up pronouns i.e. gender, number, person and case. That seems to be reasonably natural. Under this approach you would expect the anaphoric and disjoint reference conditions for PRO's to be like those holding for pronouns. So if it turns out that in a language like Italian the subject which drops is really a PRO, then it ought to have the same kind of anaphoric relations and disjoint reference conditions that the full pronoun has; roughly although there might be some difference for instance the full pronoun might be contrastive. Apart from those, they ought to behave in the same way.

The next thing is the binding theory. (see A, B and C preceding page) Again I think it's pretty minimal. Roughly speaking it says that anaphors (lexically marked anaphors) will have to be bound in their governing category that pronominals have to be free in their minimal governing category, which gives on the one hand the property of disjoint reference, on the other hand the standard properties of control for PRO. Thirdly it says that other things like real lexical NP's, which includes variables without further stipulations, since they are case-marked things, are going to be free. That's roughly the content of that, although it is formulated in such a way as to make some other things follow. What follows is that PRO is ungoverned and that trace as a variable if it is case-marked. In particular it follows that the trace of NP movement is always not case-marked. I went through the reasons for these consequences in the previous lecture. Let me try to be a little more specific about these consequences here. As far as PRO is concerned, notice that it is

-21-

still possible for it to be governed and even case-marked in the course of a derivation as long as it isn't by the time it gets to LF. There's nothing in the binding theory that stops that. That may have some bearing on the Icelandic stuff that Richie talked about. Consider further the principle that says that (_{NP^e}) is a vbl iff it's case-marked. It follows from the theory outlined that if it is case-marked it is a variable. But it does not quite follow that if it is a vbl it is case-marked. That follows for the only case I had thought of namely the case in which the trace is the trace of quantifier. That follows from the case-filter which says that the quantifier has to be case-marked because it is lexical, therefore the vbl has to be case marked. But that leaves an other possibility, namely a quantifier could move from a case-marked position and then move somewhere else and govern as a variable a non-case-marked trace. Nothing will block that. Bob May has pointed out to me that apparently there is such a case, namely the case of ambiguity involving likely and promise, compare for instance cases like some senator is likely to impress John's constituency and Some senator promised to impress John's constituency, where the former has the narrow meaning. The way I understand it would work is the former actually lowers the quantifier that now governs a variable that was never assigned case: Recall that the only reason the vbl was to me marked for case was to allow the quantifier to be marked for case, but in this case the quantifier can inherit its case from an other source.

That covers what I did last time. So let's proceed. The next thing to do is to look at what happens in the various cases, that is with PRO, lexical NP's, pronouns, anaphors, traces of movement etc.

(Remark by Richie that points out that if the definition for vbl is "trace bound by appropriate operator" then in cases like "Who was killed" and if the definition is taken literally then the trace in object position is a variable. Chomsky doesn't agree and says that the object trace is not bound by the quantifier: Only the subject is: Binding is going to be property that will hold of traces that are case-marked. I want to interpret "who was killed" like in (1):

(1) Who_x(x was killed^t)

where the structure under S is interpreted like (2):

(2) John was killed

This goes along with thinking of bound vbl's as ambiguous names. It's not that you can't make up a logic in which this assumption is not true, but I think it is quite natural to make up a logic in which it is true. The reason comes from the intuition about quantification from which the whole quantification theory springs. Think about "Tractatus": If you have a fixed finite domain then you never have quantification: You'd always use conjunctions and disjunctions. The

notion of quantification arises for domains that are either inaeirnate or somehow non specified or infinite. In a finite domain of known individuals the universal quantifiers are conjunctions and the existential quantifiers are disjunctions and variables are precisely names so the theory of quantification rather naturally generalizes from that and says that the same is true in domains in which you can't spell out the universal quantifiers as 'conjunction and so on.

(Question by May concerning the status of traces in COMP (of wh-mvt). Chomsky says that these are not variables ^{perhaps} because a vbl is a case-marked argument and bound by the appropriate operator (May's suggestion)). Chomsky notes that he is "hesitant to make the notion more precise at this point because there are still open questions, notably Taraldsen's description in his GLOW-paper where a case-marked slot could be treated as a variable but is not the position from which the wh-phrase moves. Chomsky notes that "what is required by the binding theory is that a vbl is going to be free, that is never c-commanded by anything ccindexed, which means that we are always going to have to pick as the variable the c-commanding one, that is we'll never be able to pick as a variable something which is c-commanded by an other thing with the same index")

Let's take a look at some actual cases now. First let's take a look

at PRO. We know that at LF at least PRO has to be ungoverned, so it will have to be in a position that is not governed, that is the standard such position, namely

(4) $(S, (S __ \text{to VP}))$

i.e indirect questions, promise and persuade and so on. There are circumstances in which you can't have PRO. For instance suppose you had deleted the S', for instance in a sentence like

(5) (it is certain $(S \text{ PRO to VP})$)

Here PRO is not case-marked but it is governed, namely by certain. Therefore that's impossible. In contrast (6)

(6) It is impossible $(S, (S \text{ PRO to win}))$

is all right because there's an S' and hence the PRO is not governed. The difference between (5) and (6) being that whereas impossible can take a for certain cannot. In other words this is related to what I tried to say in the previous lecture namely that S' deletion can apply if S' does not L-contain S. This might be the reflection of something else although it is hard to say exactly what it is. For instance it cannot be a semantic property of the matrix adjective since you get synonyms which behave differently, like likely and possible. In the cases where there's some overt indication of L-containment like in English for or the undeletable that then S' deletion seems to work out with L-containment, although it must be a consequence of something else because of languages like French where you can't have for. For the

rest of the time I'm not sure but that's an area where to look

for some unifying principle, although obviously there are differences between languages, so for example croire in French takes control but believe doesn't. So there's some flexibility and also some non flexibility.

(Question by Richie concerning control in Il me semble m'être trompé vs *Il semble PRO avoir pleuré (if it's a case of control). Chomsky notes that there are cases of the same type in English too, like for instance the PRO interpretation of sentences like "He said to leave". The obvious (non explanatory) answer is that S' doesn't delete in the me semble case.

Question by Richie concerning the independently motivated use of S' deletion. Chomsky notes that S' deletion is needed also because traces are going to behave like anaphors except that they are not case-marked.)

Anyway things like (7)

(7) *It seems to leave

are going to be out, i.e will have no obviative interpretation. You can't have PRO in cases like (8)

(8) *I think that a picture of PRO is on sale

for two reasons: For one thing it gets case from the inserted of and secondly it is governed by picture so that's impossible. Consider too (9):

(9) *The destruction PRO

I am assuming that the destruction of the city is just base-generated with the same structure as (9), under normal interpretation for the X-bar theory. But PRO cannot be an option here because even though destruction does not assign case it governs PRO, which is impossible. However you could have trace instead of PRO because trace only has to be non-case-marked, it can be governed which means you can get phrases like the city's destruction but not (9) with the interpretation the destruction of some unspecified cities. There are other ungoverned positions besides the one in (4) which I will come to in a minute. However before doing so let me mention that all of this is relying on assumptions that you ought to be rather suspicious about namely how exactly is government working. It works but very tenuously. So for example I am assuming that the subject of an infinitive is ungoverned which means that in (4) VP does not govern it also that V does not govern it. (VP is not a governer) that works but works so narrowly that you should worry about it. Notice too that this depends on the assumption that there's no AG(remnant) on the infinitive since I am assuming that this element is a possible governer. I don't know what happens with Portuguese with respects to that, nor do I know what happens in VSO languages since VP is playing a crucial role. I shall nevertheless continue assuming that it will somehow work.

So that takes care of the first case of PRO, namely PRO where its minimal governing category is S. What about the N case now? Here you can begin to exploit the laxity in the system I presented. We are concerned with

what happens in the determiner, that is in (10):

(10) $\text{NP} \rightarrow \text{NP} \text{---} \text{N}'$

namely in phrases like John's book. Plainly you can get lexical NP's. So we shall have to have some kind of a rule, let's call it R, that assigns possessive and let's state the rule as in (11):

(11) $\text{NP} \rightarrow (\text{NP} \text{NP+Poss}) / \text{---N}'$

So from John book you can get John's book. If Poss is introduced then it governs NP and since we take it to be a case assigner NP will get genitive case. This so far works OK. So take (12):

(12) *John likes PRO's book

that's out because PRO has to be ungoverned and it would be governed by the genitive. But now what happens when you extend the analysis in (11) to cases where the context is X' and not simply N'. So take N' and V' and in V' you obviously can get PRO. So while you can get

(13) John likes my book

you can't get (12). But now consider

(14) John prefers my writing a book

(15) John prefers his writing a book

(14) and (15) are OK but (15) has disjoint reference but in John likes his book there's no disjoint reference. At least there's a very strong tendency in that direction (which you can somehow reduce by adding himself (John prefers his writing the book himself)) in (15) that's absent in (14). Thirdly you can get PRO in these contexts so you can get

(16) John prefers writing books

What can all this mean? Assuming rule (11), we'll have to make it optional when X'=V' (i.e. the rule is $\text{NP} \rightarrow (\text{NP} \text{NP+Poss}) / \text{---X}'$) but obligatory if N', because of the fact that you can get control in sentences like (16) or He likes writing books. There will of course be lexical variation but that's the basic situation. This description doesn't yet disallow coreference in (17)

(17) John likes his writing a book

In other words it doesn't yet capture what I take to be the basic intuition namely that PRO and pronoun are in complementary distribution, or to put it in a different way that you choose not to use a pronoun if you could have used PRO, that is that there exists some kind of rule that says "Avoid pronoun". Although I don't know how to formulate this principle I think it's true and the tendency could be worded as "Get rid of pronoun if you can". Notice that we are very close to having PRO and pronoun in complementary distribution anyway, for quite different reasons, namely pronoun always has to be case marked and PRO can never be case marked. So the only case where they could be in contrast are the cases where the case marking is somehow optional. Here we have one case where the case marking is optional (John likes his writing ~~books~~ so the complementary distribution could do

-27-

but the principle that says "avoid-pronoun" tells you that in (17) there's disjoint reference since if there wasn't you'd be violating the avoid pronoun principle. Now of course there may be variations among people: some may apply a very strong form of the principle, some a weaker version. An other consequence of this is to give you some kind of recoverability of deletion rules for pronouns. The principle says "Delete them when it's recoverable, namely when you can avoid them". If this is correct then we must be looking for cases where there's an optional rule of case-assignment since this is the only structural position in which the complementary distribution could be violated. So the complementary distribution follows from deeper principles namely the case filter and the Binding theory, which requires that PRO not be case marked and that pronoun be case marked. An other case in English is the case of for+infinitives.

So in (18)

(18) John bought a book for me to read
you can substitute him for me

(19) John bought a book for him to read

In (19) the disjoint reference reading is very strongly preferred and you can get PRO:

(20) John bought a book to read

Now given the way the base works you should be able to get the three things but the avoid pronoun principle, ^{says that} since you can get PRO in (20) you're going to try not to use a proximate him, hence the disjoint reference interpretation in (19). So the principle could be reworded as "Avoid proximate Pronoun". The net effect of all this is that you appear to be getting a disjoint reference rule. But there is no disjoint reference rule. Something that further suggests that the above is correct is the following: There is, as is well known, a kind of general patterning of disjoint reference with bound anaphora. Where anaphors are necessarily bound, pronouns are necessarily disjoint. These things go together. However there are some cases where they don't, essentially the cases above. So although you do get disjoint reference you don't get reciprocals

(21) *They bought books for each other to read

Comparably though you can get

(22) They like each other's books

you can't get

(23)*They prefer each other's writing books

That follows from what I have been saying before (in (21) and (23) each other is not bound in its governing category). Notice too that it's strikingly a case where each other and disjoint reference do not pattern alike. This suggests that disjoint references in these cases is a distinct phenomenon from disjoint reference where they do pattern alike. In the above cases disjoint reference falls under the "Avoid Proximate Pronoun" principle.

(Question by J. Guéron: How far do you want to carry this analysis. For instance would you be willing to say that you can't get "Il_i veut qu'il_i vienne" because it's in violation of the "avoid pronoun principle" because of the existence of "Il veut venir". Chomsky not willing to commit himself here)

That essentially takes care of PRO for both the minimal governing category S and NP. The next thing to look at is pronoun in S and NP. Let's start with S. By the case filter they are going to have to have case. Therefore they've got to appear in case marking positions and, by C in the binding conditions (see formulation page 16 above), since they're pronominal they're going to have to be free in their minimal governing category. That gives you all the disjoint reference facts. So take

(24) John saw him

The governing category is S and it's not free under the non disjoint reference reading, so that's out under that reading. Consider further

(25) John believed that he would win

Here he is free in its governing category, namely the lower S so that's OK.

(26) John believes him to be a fool

Here you have exceptional case marking, so here the governing category is the matrix S, hence him is not free in it under the non disjoint reference reading, hence it's out. Consider the case of the SSC

(27) John expected me to like him

with John and him coindexed: The governing category is the lower S, him is free in it so (27) is OK. Same for

(28) John told me about him

where if John and him are coindexed him is not going to be free in its governing category, namely S, hence out. What about?

(29) John saw my picture of him

Remember that S and NP's are governing categories so if John and him are coindexed that's OK because the governing category for him is NP not S. What about

(30) John saw his picture of me

Same story.

(Question by Pollock concerning possibility of coreference in cases like Jean m'a parlé de lui_i. Chomsky notes that in English you can also get John_i kept it near him_i and says that no one knows how to analyse them, although you can of course make up a story, say for instance that PP is the governing category in these cases although no independently justified analysis explains why this has to be the case)

(Question by May concerning sentences like John saw a picture of him which are not very good. Chomsky notes that this is also the case if you replace him by himself. "My guess is that the rules aren't very determinate in that case because of the marking properties discussed in previous lecture: The rules are determinate for each other. If you extend them to himself, you'll never get

him (and some people like John saw a picture of himself better than John saw a picture of him). If you don't, then you can get him but not himself. It's rather indeterminate. I am assuming it's a marked phenomenon anyway)

Although I won't run through the comparable cases with each other and so on. It's all basically the same except that whereas you're going to get disjoint reference in the cases just discussed, you're going to get bound anaphora for each other because of the difference between the formulation of A and C in the Binding Theory. In other words ^{what} I am claiming is that it is a property of pronominals that leads to disjoint reference. Notice that the treatment sketched makes no use of indexing, anaphorix index, never mentions disjoint reference (the phenomenon of disjoint reference is deduced from something else, because of properties that are true of pronominals which have certain effects for PR⁰, other effects on pronouns). The basic point is that you get the SSC and the NIC straight off and the theory ^{of government} gives you disjoint reference too.

It's also better that On Binding because it gets John likes his book.

Suppose now you replace John by an appropriate quantifier phrase like everyone i.e you get

(31) Everyone him

(31) will be

(32). $\forall x (x \dots)$

and exactly the same rules will hold. If they are coindexed they've got to be in a position where coindexing is legitimate for nouns. So you get the basic principle that Jim higginbotham proposed in some of his papers, namely that a pronoun can be bound by a quantifier just in case replacing that quantifier by a name gives you a possible case of non disjoint reference. That principle exactly falls out. There are of course some cases that noone knows how to deal with, for instance sentences like the woman who every Englishman admires is his mother where "his" is within the scope of every. That should be completely impossible and becomes so if you replace is by an other verb, for instance resemble. These sentences probably have something to do with the things that Roger Higgins talked about where you also have very odd interconnections between predicate expressions in pseudo-clefts or things that have NP's in place of the subject in pseudo-clefts.

Notice too that nothing has to be said about proximate and obviative pronouns: A pronoun will just happen to be obviative if there's nothing around with which it can be coindexed, and will be proximate otherwise. Of course it won't be proximate under too local conditions but that's because of C

The next class of cases is the case-marked ones. So let's start with the case-marked non-anaphors, that is lexical items on the one hand, and variables on the other hand. Concerning lexical items there's nothing much to say: According to B a lexical item just has to be free in every governing

category (condition B should be restricted to non-pronominals). What would have been expressed in On Binding in terms of anaphoric index is now expressed in condition B. The next case is variables. Recall that if we have a case-marked trace it's a variable. So consider (33):

(33) John_i likes t_i

t_i gets case-marked, it's a variable ^{which it isn't} further more, and redundantly, it has to be free, which it isn't in (33). So you can't get movement from case-marked positions. What about the other cases of improper movement in cases like:

(34) (_S, COMP(_S ... vbl ... (_S, COMP(_S ... vbl ...)))

that is the cases that Bob May has talked about. I shall rule it out essentially on May's grounds, namely the traces being variables, namely something governed by something in COMP, has to be free, but it isn't since the lower vbl is not free in the higher S. So the solution is exactly the same as May's except that it doesn't involve anaphoric indexes and contradictory indices.

(Question by Richie How does one exclude *Who does it seem to like

with the structure (_S, who_i (_S, it seems(_S, t_i (_S, t_i to like t_i)))) Chomsky's solution is that the subject trace (which is a variable) is not case-marked. The reason for the subject trace is a variable is that it c-commands the object trace.

If you isolate the lower S you get a case of NP movement and you can say that in a string of traces it's always the c-commanding one that's going to be the variable. In fact if we were to have picked the lower trace as the variable then the structure would have been out because it then is c-commanded by the trace in subject position and therefore the variable is not free in the lower S, contrary to the requirements of B (see page 16 above).)

Now Higginbotham used the contradictory indices to provide the basic principle of pronouns that I stated before namely that a pronoun can be bound by a quantifier just in case you can replace that quantifier by a noun and get non-disjoint reference. There are in fact two basic cases. So (35)

(35) Everyone saw him

doesn't mean

(36) $\forall x (x \text{ saw } x)$

The other case is

(37) He saw everyone

which doesn't mean

(38) $\forall x (x \text{ saw } x)$

The principle is that if you replace the bound variable by a name (John) you can't get coreference in either case. Higginbotham's analysis is that in the case of (37) if you try to make he coreferential to the trace of everyone then you get

(39) $\forall x (\underline{\text{he}}_i \text{ saw } x_{\{i\}})$

which is supposed to be absurd. Therefore it's out. Notice that on the analysis

suggested above we don't need any of this. (35) is dealt with by C. In (37) the object trace is a variable, it should be free but isn't hence that reading is out. This simply is the result of regarding a noun as a bound variable. Let me make two observations concerning the comparison between the two theories. First of all the one that doesn't require the use of anaphoric indices is obviously the better one, simply because it doesn't have the whole mechanism of anaphoric indices, the notion free_i, the notion designated index and so on. Furthermore, although the facts are a little slippery I think there's some empirical evidence bearing on the choice to be made: If one says something like:

(38) (NP_{i{i}})

is absurd, the question is why is it absurd? Presumably one wishes to say that that's absurd because there's a principle that says that in some sense variables or their instantiation can't be coreferential with names that have the same index. So it's a principle that says that if you have

(39) Carter voted for everyone on the democratic list
Carter can't be on the democratic list. But that's not true. Similarly take

(40) Who did Carter vote for?

A perfectly fine answer to (40) is Carter. The same is true for (41):

(41) Who voted for Carter

where the answer could be Carter again. Similarly for existential ones. Suppose you say

(41) Someone in Plains voted for Carter

and it turns out that it's Carter, that doesn't make (41) false. Quite generally it seems to be a property of variables that they do not enter into disjoint reference, that is, at least, their instantiation need not be disjoint in reference from the thing appropriately related to the variable. That is, the principle on which we base the absurdity of (38) is false. Notice that on the theory I am now talking about nothing is said about any kind of disjoint reference in these cases. The only thing it says is about coindexing, namely that Everyone saw him cannot mean $\forall x x \text{ saw } x$, and that's the one fact that is clearly true. The others have to do with pragmatic phenomena and how you decide what your universal category is and so on. So I think that the present theory is both simpler and factually better.

Now let's take the case of strong cross-over, so cases like

(42) Who did he say Mary kissed

which will be represented as

(43) For which x he said Mary kissed x

Again for all these cases the facts are a consequence of B. and nothing further needs to be said. In other words the "cross-over" phenomenon doesn't exist

Notice that it follows from what we've said so far that variables are immune from the SSC and the NIC. Notice too that the SSC and the NIC are just theorems here, and they're not going to apply to variables. That's correct in principle, since variables are names.

Of course if we say that, then we have to solve the problem of why something like the NIC appears to hold for variables in cases like *who did he say that had won. What I shall suggest is that the NIC in that case is a spurious generalisation.

Now the next thing to worry about is trace of NP. If you think about the structure of the system it's obvious why it is going to work the way we want, namely because the notion "governor" is broader than case-marking. Some governors are case markers others not. It follows then that the trace of NP will have the same properties as each other. So for concreteness consider

(45) John was killed t

The trace is not case-marked, it is bound in the governing category (S) so OK. Consider further

(46) John was believed t to be a fool

The governing category here is the higher S, the trace is governed in it so OK

(47)*John was believed(_St was killed))

here the governing category is the lower S (because tense is assigning case, which is redundantly incompatible with what we saw concerning trace of NP movt) it is not bound in it, hence out. Consider

(48) John was believed (_SNP to have killed t)

A case of the SSC that follows from the fact that the trace isn't bound in its governing category (i.e the lower S in which it received case) (question by May on things like * John believes to be incompetent. Chomsky says that that's out because the trace is case-marked. So despite the fact that it's bound in its governing category, it will fall under B. i.e should be free in its governing category, which it isn't, hence out. Rizzi notes that for this solution to work it's got to be the case that S' deletion obligatory applies in this case. Chomsky notes that if S' deletion didn't apply the sentence would still be out for other reasons, in fact for the same reasons as the reasons that prevent *John was tried to leave namely these structures have S' and S. It's going to turn out, when all the principles are in place that there's a theorem that says S' is an absolute boundary for movement. That's an obvious desideratum and it is in fact going to follow from the Residue of the NIC)

Take a look at raising now. Consider

(49) John seems t to VP

(49) is OK because the trace is governed (though not case-marked) and is

the subject too far, as in

(50)*John_i expected me to seem t_i to be happy

(50) would be out because it would not be bound in the intermediate S, which is required by A. If raising applied to the subject of a tensed embedded clause the structure would be out because of B. Similarly for raising applying to the object of a verb. I haven't talked about the case-marked anaphors, namely each other. It's plain that the main cases are going to work exactly as expected, that is you'll get the SSC and the NIC, because of the properties of governors. There are however some other things to be said about reciprocals, and I don't understand them, I'll just say them, namely reciprocals seem to be much more natural in what is sort of an object position, so things like (51)

(51) They left with each other

don't seem too good to me, although their meaning is fine. Similarly

(52) They are easy for each other to talk to

is rather weird, at best, meaning each is easy for the other to talk to, Notice crucially that here each other is up on top. Sometimes with adjectives and for it's not too bad, like in

(53) they are sorry for each other

my intuition being that sorry for is reanalysed somehow and each other is not the object of for really. So it seems to me that we're missing something about reciprocals. They look as if they were subcategorised by the verb and the cases where they seem natural after prepositions like in they spoke to each other are probably cases of reanalysis. What I would really like to believe is that the possibility of taking a reciprocal or a reflexive is really mentioned in the verb, like for the reflexive it is, I suppose, in Romance. So let's imagine that verbs take reflexive clitics which are realized as clitics in Romance and are just varied in English and that something like that also works for reciprocals. This relates to some of the points that Richie was making yesterday. It's also true, as he pointed out, that in cases like they were happy for each other to leave. My feeling is that this doesn't have much to do with this theory because I think that that's also true when each other is up above as in the easy cases. I say this very tentatively because I really don't have a theory about what's going on. If it's true that reflexives and reciprocals are properties of the verb it would follow that you can't get them in subject position.

(Richie notes that if this is true the contrast between reciprocals and possessive in cases like (51) is better expressed. Chomsky thinks that the possessive is marked in English anyway. Richie points out that in lots of languages the reflexives are much freer than in English, in Scandinavian they occur down in lower sentences. Chomsky notes that in that case they are different kinds

-32-

of items and although they are not as "free" as pronouns they are not "anaphors" in the sense his theory gives to the term. Van Riemsdijk notes that it wouldn't be true to say that reciprocals can only occur in those positions where you get reanalysis due to things like "They told Mary about each other". Chomsky agrees.)

Next thing is raising over adjectives. If the adjective is an S' deleter, so for instance

(54) e is certain (_S NP to VP)

in (54) plainly the NP can't stay there because of the case filter, however it can move to e, yielding

(55) NP_i is certain (_S t_i to VP)

Here everything is fine, the trace is governed by certain, the higher S is its governing category, it is not case marked, hence is not a variable, and therefore has all the right properties. The NP couldn't have moved if the embedded S had been tensed, it couldn't have moved from the object position it couldn't have moved from for for the same reasons as each other.

This takes care of the main story. Let me make a few side remarks about other cases. Notice that you can sort of get a version of A/A out of this for free, for some cases at least. Suppose you had things like

(56) NP V (_{NP} NP's N)

if you moved the inner NP, its trace would be free in its governing category. If you didn't have the 's the trace would be ungoverned and for other reasons you can't have that either. Notice however that in cases like

(57) NP V (_{NP} the destruction Rome)

you can get (58)

(58) NP V (_{NP} Rome_i destruction t_i)

but not

(59) Rome_i V (_{NP} the destruction t_i)

because the trace is governed by destruction, hence has to be bound in the inner NP (by A.) therefore can't move outside of it. Note that if the thing that moves leaves a variable, then it shouldn't abide by these conditions and I think that Richie's Tous-mvt in the paper for the Revista (i.e L-Tous, CL-PL, Binding and Subjacency) has the latter character.

I'll be assuming that what I have said so far is correct and I'll start from there. From the Binding theory it follows that variables are free in every governing category. From that it follows that they are also free from any effect of the SSC or the NIC. Notice too that the NIC and the SSC remain true, they are sort of theorems of the present theory, I say sort of because it doesn't work exactly the same but there's large overlap and so they remain true for a considerable mass of cases. I think that that's a desirable result and ought to be a desideratum for a theory and my reasons have to do, pretheoretically, with the feeling that the deeper meaning of quantification really relates bound variables and names in two fundamental respects, first that names are instantiations of variables, secondly that quantification as it historically developed, and as logically it ought to, was related to conjunction and disjunction so in a way it's kind of an idealisation from the constructive case where you know the elements to a non constructive case where you don't know the elements. But then, as an idealisation, it ought to share the fundamental property, namely that in fact variables are just unspecified names. So it seems to me to be a plausible conclusion on the basis of the meaning of quantification. Anyway it comes out. There is a number of reasons as to why one wants to lead to this. One is the fact that, as first pointed out by Rizzi, wh-traces at least don't observe the SSC. The second point which came out in the Freiden and Lasnik paper is that they don't observe either the SSC or the NIC with respect to cross-over. A third fact has to do with Bob May's explanation, which I just carried over, for the improper movement cases; i.e. if you have improper movement where the most deeply embedded trace is the trace of a subject you want it to be argument bound by the top one, that is not to be subject to the NIC. So these three things converge to what I take to be a desideratum, namely that bound variables are just unspecified names. Now there is a problem, a problem that's been talked about a lot, namely that you don't get

(1) *Wh_i... (S_i nom ...)

same with indirect questions and so on and so forth. It is very clear that those are no good. And what makes them look no good looks awfully much like the NIC, but, as I have been saying, for these reasons and some others, that was a spurious generalisation. It's really something else that apparently converges with the NIC in this particular case. The NIC, now absorbed in the Binding theory says things about anaphors, NP traces, pronominals and lexical items. But the only thing it says about bound variables is that they are free. Similarly some other condition will say similar things about bound variables and some of the things it will say look like what the

the other theory says about bound traces, but that will come from an other source, I suggest. Let me look at these points a little more. There is, as these examples indicate, something like a contradiction between the assumption that wh-traces don't observe the SSC, don't observe the SSC nor the NIC with respect to cross-over and with respect to improper movement one the one hand, and that it is the NIC that is involved in the explanation for why (1) is out. Notice that it is only nearly a contradiction. So for example suppose that you take the On Binding theory for indexing (anaphoric indexes and so on) then one might ^{find} a way out of this apparent contradiction by saying that the NIC holds for referential indices, therefore blocks movement, but does not hold for anaphoric indices, therefore permits all the other stuff. As far as I can see, although I can't imagine that it's correct, that's a coherent answer. In fact one might imagine other ways out. So as Richie pointed out yesterday we might assume that the NIC just happens to be vacuous for some other reason in the movement cases, because of something else. That's also a possibility. So contrast between (1) and what is otherwise assumed in the case of cross-over, improper movement etc is not a real contradiction although it is a *prima facia* contradiction. Nevertheless my own feeling is that it is a real contradiction. Other questions of course arise about all this. If in fact (1) had something to do with the NIC we'd want to know why we have the asymmetry, that is why the NIC but not the SSC holds for wh-traces, why not the other way round, that is why we don't find some other language where it's the SSC that works but not the NIC. Of course the problem doesn't arise if the principles that account for the two series of contradictory facts are unrelated. there's still a problem, namely what's going on in (1), but that has nothing to do with symmetries or asymmetries in the Binding theory. In fact a sort of related conceptual question is this: As is well known, there are languages, Italian being one, in which in some respect to be made precise, the NIC is vacated, that is the cluster of properties that go along with case in (1) don't hold in Italian, although they hold in French or English. So there's some kind of way that languages have of exempting themselves from some aspects of the NIC. Then the question arises as to why some languages don't have a way of exempting themselves from, say, the SSC. That is, why don't we find a language which has a parallel cluster of properties with respect to the SSC. Assuming that we don't, it would make it very plausible to say that (1) has nothing to do with the NIC. So there are many considerations, the main one to my mind being the meaning of quantification, that suggest that (1) has nothing to do with the NIC, though there may be other deeper reasons for why you have conditions that do for bound-variables what the NIC does for other things.

So, then, I am assuming a separate phenomenon and let's call (1)

the Residue of the NIC (RESNIC). The next question is what are some of its properties. You can get some kind of insight about that by looking at the languages where it doesn't hold, so Italian for instance, and see what is not happening that's happening elsewhere. The first point to make about Italian is that the actual Binding theory seems to hold without exception in Italian, so the actual NIC holds for disjoint reference exactly as elsewhere and similarly for NP traces the same thing holds as elsewhere, so you don't get

(2) NP_i sembra che t_i .. tensed V

So it's not the case that the NIC is vacated in Italian, nor is it the case that null anaphors aren't subject to NIC, they are. So that's not what's going on. That is, as far as I know, in every respect so far discussed, i.e. null anaphors, disjoint reference, whatever else, Italian doesn't differ from other languages. What does not go the same way in Italian is wh-traces — There are other things that differ of course, subjectless sentences for instance, I'll come back to those — but the crucial difference is wh-traces. So one thing this RESNIC holds for is wh-traces. We want to see what else it does not hold for, but we know we can't be too broad. We couldn't just say the NIC is null, or we couldn't just say it's null for null anaphors, so we've got to be rather circumspect. Incidentally notice that, just to take a sharp contrast, if you have a control verb in a structure like (3)

(3) crede che (tensed V)

(3) will be OK, that is, he believes that will come, but, of course, here I am assuming PRO which, remember, has features etc. is a pronoun that doesn't have to have a phonetic matrix. So (2) and (3) are in direct contrast. I'll come back to the meaning of that later on, and to the fact that here PRO is case-marked. Let's look a little more sharply at what's covered by RESNIC. The first question that arises is the question of variables. This is very complicated. So question: Does this RESNIC, this cluster of properties, long movement, wide scope etc, hold for variables? There are a number of problems, the first one having to do with languages like Italian. If the RESNIC extends to variables and doesn't hold for Italian, for whatever reasons, then Italian variables in quantified expressions ought to have quite different properties from variables in French. I am pretty sure that's not true. I am going ultimately to assume that Richie's facts require the RESNIC to be a LF property, but I want to mention the problems that will very much have to be solved if that's accepted. One is the problem just mentioned, namely that Italian behaves as all other languages with regards to quantifiers. If RESNIC extends to variables it's by no means obvious why that should be true. A second kind of problem has to do with Focus. There are lots of ways to think about focus, but one way that seems to have nice properties is to assume that focus is a manner of forming indefinite descriptions, so

(4) JOHN saw Bill

would be represented as (5):

(5) The x such that x saw Bill = John

and that fits in nicely with weak cross over and has lots of other nice properties. If that's the way to handle focus then we immediately have a problem because focus is going to leave a variable in the subject position in embedded sentences, obviously. So the representation of (6) will be (7)

(6) John said that BILL liked Mary

(7) the x such that John said that x like Mary is Bill and (7) will enter the cross-over the way it ought to, but the variable is going to be different from the wh-trace. So that's a second class of problem. The third class of problems has to do with the matter of broad-scope interpretation quantifiers, and here the facts get fairly hairy. The points to look at are going to have to do with things like

(8) John expected that {some of the talks} would be too difficult for x and to make the judgments a little sharper you can put an other quantifier in the subject of the matrix:

(9) everyone expected that {certain talks} {some of the talks}.....

so you get an order of quantifiers problem. A lot of people believe that they get wide scope interpretation here, especially with certain. Maybe they are wrong and it's narrow scope interpretation, but if you do get wide scope interpretation, then it will mean that variables will be left behind in positions where you can't get wh-traces. The same problem arises if we invert the order of quantifiers (someone expected that every talk would be too difficult). Can that mean "for every x there is y such that y expected that x would be too difficult". I don't know, but it doesn't sound wild. If it is true then it's a problem for extending the theory to variables. Actually there's a narrower, more technical version of this problem that arises just in simple cases like (10):

(10) Everyone loves someone

I don't think that in (10) there's any doubt: Every can have wide scope hence you get (11)

(11) $(\forall x (\exists y (x \text{ loves } y)))$

but in (11) the relation between $\forall x$ and x is not a relation of local c-command, that is $\forall x$ is not locally governing x, that is, what is governing x is too far away in LF. This problem is a technical, not a conceptual, problem, unlike the other three.

Then of course the major problem that arises is the problem of so called broad scope quantifiers, any being the obvious case in English. As is well known, in many cases you can interpret the broad-scope quantifiers

either as universals or as existentials. So I didn't see anyone could be represented as $\forall x$ I didn't see x or as $\exists x$ such that I saw x . That leaves you a lot of options. However sometimes it's rather tricky to see how these options are going to be exercised, the main class is where you get quantifier interaction and the obvious cases to investigate are indirect questions. So consider

(12) I wonder how anyone understood that talk

Here there's quantifier interaction in a sense. That is (13) should be out

(13) I wonder how $\exists x$ such that x understood that talk

Notice too here that the broad scope interpretation at least gives you the right logic:

(14) $\forall x$ I wonder how x understood that talk

I think that one might argue that something is lost in the translation in (14) although all the inferences come out straight, as far as I can see. Anyway (14) is a correct interpretation of (12), whatever it may loose in flavor, but it leaves a nominative variable in the kind of position that ought to be impossible. These problems become magnified the more you think about examples of that kind. So for example let's take other examples that illustrate the interaction of quantifiers, like (15):

(15) Noone knows how any of us did in the exam

The obvious logic of (15) is

(16) $\forall x \sim x$ knows how x did in the exam, $x =$ one of us

But in (16) there's a nominative variable in a position that contrasts directly with the position of a wh-variable, since the latter would not be allowed.

The same problem arises in the case of (17):

(17) They couldn't tell me how any of us did in the exam

which seems to imply, by universal instantiation, they couldn't tell me how I did in the exam; which again would make sense if you have wide scope. Robert May is suggesting that this may also be true for the non-wide-scope quantifiers (like every for instance). If that's true, then the problem is even worse, although they don't sound as obviously right as the any examples, so for instance I think that they couldn't tell us how every one did on the exam can be interpreted with narrow scope but any in (17) cannot. That is as a universal any never permits the internal scope in structures like (17) whereas if you have every you can have both narrow scope and universal quantification. Therefore we are left with two options: One is that any is a wide scope universal, as in (17). The second option is that any is a narrow scope existential. Of course there might be a totally different way of looking at the problem, but if that's the right option, then these two exhaust the possibilities, and the first option is the correct one in the case of (17) although, as I said, the translation loses something.

-30-

Then there are all Richie's cases which go the other way, that is they all indicate that variables are treated like wh-traces, multiple interrogation and negation.

(Richie notes at this point that in those two cases you have indication of scope in the syntax)

So these are the main problems that I can see in an attempt to extend the RESNIC to variables which means in effect letting it apply at LP. Of course if you define it on the phonology, then you have all the problems that Richie mentions. So of course there's a contradiction and let's hold these problems in obeiance. Now let's look at the next problem that is the problem of determining the phenomena that the RESNIC will have to deal with. Work on Romance in the last couple of years has brought up a number of phenomena that apparently fall together. One of them is the question of missing subjects. So some languages have their objects missing, like Italian. The second type of problem is the free stylistic inversion which you get in Italian pretty freely, in matrix sentences for instance. In French you sort of get the phenomenon by only in the special constructions that Richie and Jean-Yves were discussing. The third difference is the long wh-movement, which you can get out of subjects in Italian, but not in French and English. The fourth problem which Tarald discussed recently in a paper has to do with a similar phenomenon but which can't be wh-movement, it's got to be a resumptive pronoun strategy, an example he gave was something like

(18) *ecco la ragazza (che (mi domando (chi e tensed V))*

where the "e" cannot be the trace of wh-movement, because of adjacency. So Tarald interpreted the "e" as base-generated, so let's call that a resumptive pronoun strategy that allows a null resumptive pronoun in Italian, which is impossible in French and English, in the dialects that have resumptive pronouns. The fifth phenomenon is that the "that-t" filter, or whatever its proper formulation is, doesn't hold for Italian but holds for English or French, as Perlmutter pointed out. Basically what seems to be the case is that either a language has all these properties or none. Therefore we have to assume that there's some kind of language parameter somewhere that's fixed one way or another, and there ought to be enough to determine whether a language has one of these things or none of them. Of course it's more complicated but let's start with that. I am going to assume that the parameter is of the kind that Taraldsen talked about in this paper. His idea is that this has to do with agreement. The intuitive idea is that where you have overt agreement, you can drop the subject, intuitively something like recoverability of deletion. So Italian has a kind of agreement that French doesn't have. Remember I am assuming a structure like (19):

(19)a NP INF VP

b INF, + Tense, AG (person number gender) . . .

As you can see, there's no difference between AG and PRO:AG might as well be called PRO. Now we want to say that the difference between French and Italian is as small as possible so let's say that they have the same base structures, that is they both have AG, but that Italian has a further property that French doesn't have, namely that Italian has indices on the AG. That seems like the smallest property one can think of. So there's this parameter that says that in a language AG may or may not coindex, that is, has an index since under free indexation it will coindex. This is at the level of execution, and there are other ways, but let's try that one. Now let's formulate the RESNIC as in

(20) (_{NP^e}) cannot be free in S'.

nominative

Now, given (20), coindexing with AG will get the first four facts: Since coindexation can work in Italian but not French, then long wh-movement will be OK in Italian but not in French, (_{NP^e}) will not be free in Italian but will be in French, hence the pro-drop phenomena and the stylistic inversion phenomena, same again for the resumptive pronoun case. The only problem that comes up then is the "that-t" filter. How we implement this depends on a decision to regard the RESNIC as being on the phonological side or the LF side. Let's try it both ways. Suppose RESNIC is in the phonology, is a filter and hence applies after deletion. Then the obvious thing to relate it to, as David Pesetsky pointed out, is the doubly filled COMP filter:

(21) *(_{COMP^{αβ}})

Suppose now that trace is one of the possible choices of α , so if you have both the trace and that it's out, by (21). An other possibility is just the trace and no that, which of course is the one that's OK. The next is no trace and that. That's bad (who do you think that left) because of (20) since the nominative subject is now free in S': Just as the AG will not bind the subject in French, similarly, in English it won't, and if there's no trace in COMP, then the COMP won't bind it either. so it's out. So (21) gives you the that-trace filter free and makes some predictions, namely that languages that allow doubly filled COMP's should permit a violation of (20). Note that so far (20) only distinguishes movement from subject and object by stipulation, by mentioning a nominative trace, which can't be right; But let's start with (20), I'll come to an other formulation later on. The stylistic inversion facts in French will have to be treated in the way Kayne and Pollock treated them, namely there's a rule in French that moves a Q into a position that's controlling the trace. So this thing seems to be an extremely simple way of bringing these things together, so far assuming (20) to be on the left in the phonology. Of course, if we assume that, then

we have no treatment for handling Richie's facts. But then we don't have the other problems. So now let's try it the other way, ... (20) applies on the right, at LF for instance. Now of course we shall be in a position to deal with Richie's facts, but the other facts will become problems: That's a problem for future research, and I have no answer. Let's try that option anyway. Essentially we get the same argument. First of all as far as the first four properties are concerned there's nothing to say since choice of right or left is completely irrelevant. The only question is the that-trace filter, and it comes out in exactly the same way. Remember that the rule for COMP is

$$(22) \text{COMP} \rightarrow \{\text{for, that, +WH}\}$$

(22) like anything else is optional, so it can not be chosen. I am assuming too that there's some kind of convention that has the effect in (23):

$$(23) (\text{NP } e) \rightarrow \emptyset \text{ (optional)}$$

(23) will probably be part of the "avoid-pronoun-principle". Note that (23) can't apply if trace has case. So an uncase-marked trace drops. That's Richie's suggestion in the Revista paper. Recall too that I was assuming that when you get repeated wh-movement, the convention for assigning case is minimal, that is the wh-word inherits the case but the traces don't. So those are the assumptions and let's try to think of the possible examples we have. (23) applies before the two branches of the grammar separate, at S-structure or earlier, it could be a convention that applies whenever there's a non case marked trace. Let's assume it applies at S-structure. Now consider

$$(24) \text{John thinks that Bill saw something}$$

You can get

$$(25) \text{What does John think (t that) Bill saw t}$$

or

$$(26) \text{What does John think (t) Bill saw t}$$

(25) is the base possibility that picked the that, (26) the one where it didn't. Then there are two other possibilities to worry about, namely movement out of subject. You can have either

$$(27) \text{who does John think (t that) t saw Bill}$$

or

$$(28) \text{who does John think (t) t saw Bill}$$

In (25) one possibility is that the trace deletes by (23). If it deletes the sentence is OK. If it doesn't delete, then it's out by the doubly filled COMP filter but you never know it. In (26) applying or not applying (23) will always yield a good sentence. Consider (27). Suppose the trace is deleted then (27) is out by RESNIC, since the nominative trace is not bound in S'. If it does not go to zero, then it's still bad by (21), the doubly filled COMP filter. In (28) if (23) applies then out by RESNIC. If the trace does not delete then (28) is OK. So here again we get the that-trace filter with no

extra machinery at all, so that you deduce the that-trace filter, strictly. Notice that this result is completely neutral with respect to choice of left branch or right branch for the RESNIC, although I believe with Richie that it is on the right. So anyway we get the correct clustering of properties without any extra apparatus. The only thing we need is some parameter which we need anyway since Italian is different from French. The parameter I am suggesting, more or less following Taraldsen, is that whether the AG is indexed hence can properly bind.

(Richie notes that there are cases where the subject is empty even though the verb form is not inflected at all, as in gerundive structures where you can say in Italian "having rained, John went to the store". Rizzi says that gerunds contrast with infinitives since you can't get those in infinitives. Chomsky says that he is forced to the conclusion that gerund can bind the empty subject)

So apart from that, it seems that everything holds together very nicely. Before talking about the proper formulation of RESNIC, let me talk about some other interesting ideas of Richie's again. The first class of examples are the differences that he pointed out between transitive verbs on the one hand and intransitive verbs and adjectives on the other as matrix clitics depending on the issue. The examples are:

- (29) John who it is essential she talk to t
- (30) John who it is essential t talk to her
- (31) John who it is essential that t talk to her
- (32) John who it is essential that she talk to t

That's one set of examples, the other is:

- (33) French, which it doesn't seem Mary knows at all well
- (34) Mary, who it doesn't seem t knows French at all well
- (35) Mary, who it doesn't seem that t knows French at all well
- (36) French, which it doesn't seem that Mary knows at all well

Richie's judgements are that (29) is OK but (30)-(31) are bad. Similarly that (33) is OK and that (34)-(35) are bad. My own judgements differ. For me (29) is slightly bad (for me there's a very strong tendency to use that in those cases, as in (32)). (30) for me is perhaps a little more questionable than (29). But for me (31) is wildly hopeless. That is, I think the real difference is between (29) and (30) on the one hand, and (31) on the other. Similarly for the second set of examples. (33) is OK, (34) is questionable (so I agree with Richie's distinction, although this is subtle and people may differ) but again (35) is completely hopeless. So my feeling is that there are two things going on here: One is the that-trace filter, and that's what is distinguishing (29)-(30) from (31) and (33)-(34) from (35). And there's an other thing going on making the other distinction, which may very well be

related to the that-trace filter. We already know how to account for the first difference. As for the second one we can borrow Richie's idea. So suppose it's true that there's a marginal process of case marking that goes in COMP. Suppose in fact that verb can assign case in structures like (28), across S'. What's going to happen in the case of sentences (25)...(28)? Nothing for the first three. As for (28) recall that one derivation lead to an ill-formed structure. But now this is no longer the case since case making into the NP in COMP will make it impossible for (23) to apply. So we get more ways for the sentence to come out right. My suggestion, then, is that that is what accounts for the marginal differences. Note, too, in that respect that when you start changing the matrix verbs or adjectives you sometimes get very different judgements. As Richie pointed out in his paper there are some adjectives which do sound OK, so for instance consider

- (37) John, I am not sure (certain, convinced) will ever win the elections

Notice that here the subject trace is in a non-case-making position and yet they are perfectly OK. In fact, to my ear they are a little better than the one that should be better, like

- (38) John, I am not sure Bill will even like

Similarly if you have NP's, that is, things that will not assign case, you can nevertheless have things like

- (39) John, I don't have the impression will like this book

(39) to me is indistinguishable from (40):

- (40) John, I don't have the impression Bill will like

Similarly if you replace seem by appear you get things like

(41) Carter, who it appears is going to be reelected president which sounds OK to me. The point I am trying to make is not so much that the judgements disappear but how they disappear: One thing that doesn't disappear when you change verbs and adjectives is that sentences with that remain completely impossible.

(Richie at this point says "the way I was looking at it, which unfortunately requires a derivative theory of grammaticality judgements, is that the grammar says the fundamental cut off point is between, say, (29) and (30)-(31). Then things like (30) are retrieved because they interfere with parentheticals. There's one argument that suggests that that might be right, namely that you can mimic these judgements in a construction such that there's no possibility of interference with parentheticals and in French the judgements are sharper (que/qui phenomena)")

Anyway, at least on the English data, the fundamental difference is between (31) vs (30)-(29). There's a perfectly plausible explanation as to why that should be the case: Borrowing Richie's idea about case marking into COMP but putting it in those terms: Somehow you can make a mental calculation so that the sentence comes out right and

if there are more of them the sentence is a little better. That could vary with the Verb and adjective chosen etc. Anyway I'll be assuming that from now on. If that's true, we can account for Richie's data by the minor modification I just discussed.

An other question has to do with an other problem of the that-trace filter which I haven't mentioned, namely that in English and the other languages that have it there seems to be two structures in which it's violated, one is relative clauses, as in

(42) the book that is on the table
and clefts

(43) It was the book that was on the table

The way that's accounted for in F&C is just by some descriptive comment. The problem is even worse in that the same violation occurs if the relative clause extraposes, so

(44) The book appeared that was on the table
or so is still OK. Richie pointed out somewhere that if you have the filter on the right then of course it will work because in (42) the wh-word will still be there. That takes care of the whole thing. Now if the RESNIC is a property of LF or is anyway independant of the phonology. This, to my mind, is a very convincing argument about the LF character of RESNIC, perhaps the most convincing. It's not that you can't get the facts if you stipulate that it's on the left: You can do it as in F&C for example, but there's no question that this is a better solution because here you don't have any device. This decision leads to the conclusion that the problematic cases are those cases where the NIC doesn't appear to be violated in cases like Focus, broad-scope quantification and so on. So let's assume that RESNIC operates at LF. If so then we could assume that Focus applies after the RESNIC is checked; as for the quantification problem, it remains to be worked out how it relates to quantifier raising.

So now we are down to the next natural question, namely, what is the status of this thing I am calling the RESNIC. There are three major problems that arise with the formulation given above. The first one is what is the notion "free" that it makes use of, what does "free" mean exactly? The second question is why the asymmetry, that is, why is the RESNIC talking about nominative and not something else? The third question is that, in some sense, this has something to do with recoverability of deletion (so for instance recall what was said about AG above), that is, one might hope that the principle of recoverability of deletion, at least for pronouns, would follow from the proper formulation of the RESNIC. Some such connection would be nice.

Let's look at the first problem, i.e the meaning of the notion free as used in (20)? If you think of all the relevant examples, their structure is

essentially as in (45):

$$(45) (S, (COMP))(S (NP^e) AG \dots)$$

the AG(reement) element may or may not have an index, depending on the language. Now if you have a coindexing relationship between COMP and empty NP or between it and AG, you have a close relationship between COMP and subject or between subject and AG and everything is OK. If you don't, then the structure is out. So the notion free as used in (20) has to do with local control and we need some notion of being controlled locally. Notice that we have the notion already, namely government. In fact government is just the relation of control between two locally related elements. So let's assume that the answer to the first question is something like government. Of course we need to be a little more precise, but notice that it has the right properties, that is there are things that do govern the empty subject, for example for. Similarly if the theory of exceptional case marking in the case of believe is correct there are verbs that also govern the empty subject. So we are just extending the structure of governors for that case. Similarly tenses and AG govern the NP. So then let's try that and let's adopt (46) as a first approximation:

$$(46) (NP_{nom}^e) \text{ must be governed}$$

Now let's turn to the second question, namely the question of the asymmetry. Why does it hold of subjects and not of objects? Now the answer is transparent: We can relate that to the asymmetry of governors. In short everything in the object is governed by the V. In fact there is really no asymmetry. We can reformulate (46) as

$$(47) (NP^e) \text{ must be governed}$$

and it will turn out automatically that NP's on the right will always be governed, therefore they can wh-move. The reason why it had to do with nominative is because that's the only place where there are options about government. Elsewhere there are no options. We of course have to work on the notion governor but that's essentially what's happening, that is if we can work out some proper notion of governor the asymmetry will fall out. In short there are two very plausible answers to the first two questions. The notion free is really the notion government and why the asymmetry, well, no asymmetry. So now we can give (47) a new name, namely the Empty Category Principle, to show that it has nothing to do with nominative and so on.

Now the Empty Category Principle is a very plausible principle.

Now we've got to worry about making it work properly. Let me introduce the notion Proper Government:

$$(48) \alpha \text{ properly governs } \beta \text{ iff } \alpha \text{ governs } \beta \text{ and}$$

$$\text{a)} \alpha = [\pm N, \pm V]$$

$$\text{b)} \alpha \text{ is coindexed with } \beta.$$

-42-

so if something is both governing and coindexing I'll call it properly governing. Now that includes all the old cases of government except one, tense government which is excluded by (48). It includes the following two cases:

- (49) (_{COMP_i} (S (NP_i e)))
(50) (S (NP_i e) AG_i)

(49) describes a proper government situation. Notice that this wasn't government before. Let's also assume, as I have been saying all along, more or less, that it was wrong to talk about nominative assignment as tense government, and let's assume that that's a structural case in tensed sentences. If we do that then (48) is just the notion Government. If that approach is correct, then it was a spurious generalisation to say that tense assigns nominative for the same reason that V assigns objective. What is really happening is that there are structurally assigned cases, like anything in the VP becomes objective and everything that's not in the VP when you've got tense gets nominative, and then there are inherently assigned tenses, in the Base; as Richie suggested, based on strict subcategorisation. If there's some c-command requirement between tense and subject for nominative assignment then we could say that the structural case nominative is assigned when the subject is c-commanded by tense. That doesn't do any worse than the other theory, because notice that the Verb-case assignment didn't really have anything to do with c-command. It had to do with strict-subcategorisation, which just happens to look like c-command. But the real point about case assignment is that it really goes along with subcategorisation, and subjects aren't subcategorized by the V. We could make them subcategorised by tense. But then we don't have to make tense a governor, that is we gain nothing by doing so. In fact we'd lose something because now the governors are a much more plausible set. This requires a slight readjustment of the definition of the notion "governing category" in ways perfectly obvious.

Just a few comments. Notice that (47) doesn't apply to PRO, that is, there's nothing in here that says that PRO has to be properly governed. So things like

(51) I don't know what PRO to do
are perfectly OK although PRO is not governed. Notice too that, so far, the whole discussion has been based on wh-traces. But there's a completely independent argument leading to the same conclusion based on NP movement: One of the problems that we have to handle somewhere is the fact that in sentences like

- (52) (John tried (S, (S PRO to leave)))

we can't have trace instead of PRO. The fact that we must have PRO is already taken care of: That position is not governed so if we had (NP e) that would be out by the empty category principle. But why can't we have trace, that is why

isn't try a raising verb. Well we know that it isn't a raising verb from the facts: If it was ~~we'd~~ get idiom chunks ~~and~~ other stuff that we don't get. Also you can't get passive so (53) is out

(53) *John was tried ($S, (S_t \text{ to VP})$)

Note that we now have the answer: If we had trace it would have to be governed, but it couldn't be, so (53) is out by the Empty Category Principle.

Let's look at the for cases now. Why can't you get

(54) *John was prefered (for (t to win))

or

(55) *John prefers (for (t to win))

Now there are two possibilities and four cases:

(56) John prefer ($S, (S_{NP} \text{ to VP})$)

in (56) for may or may not be there in the base. Let's assume that it is there. Then it assigns case. If so the trace has to be a variable but of course it cannot be since it's not bound by an appropriate operator. Besides, by P, it would have to be free in S' . If for isn't there in the base, then the trace falls under the Empty Category Principle, i.e it reduces to the try-case. So either way you can't raise with passive or non passive in the for infinitives. So in fact we have a theorem that S' is an absolute boundary for movement. (Question by Richie How does one avoid the movement through COMP analysis for try: If that was possible then the subject trace would be bound. Chomsky says this reduces to an other question that hasn't yet been solved, namely what's going to block improper movement from COMP to sentences: It's ruled out for variables but hasn't been ruled out for those cases. In other words whatever rules out improper movement from subject or object to COMP and then back into the sentence will solve the problem)

Now let's just see what happens with adjectives. So take

(57) It was certain ($S_{NP} \text{ to VP}$)

(57) is an S' deletion case, I assume for principled reasons, L-containment or something else. That means that you've got to have movement: If the NP didn't move it wouldn't get case. Why can't you have PRO? That is why can't you have

(58) It was certain ($S_{PRO} \text{ to leave}$)

meaning it was certain for some unspecified person to leave, like in

(59) It was possible to leave

The reason I think is that in possible there's no government but in certain there is government. Remember that PRO has to be ungoverned. That's a theorem. Notice too that it's going to end up as a consequence that it will be a subject only if you have a lexically filled COMP, that or for.

Let me try to say something about recoverability. I have a possible

approach to this. The whole thing depends on questions about stylistic inversion, clitics, things of that kind that I know little about. I'll try and sketch a way of relating all this to recoverability. OK so let me make some general points; we want French, Italian and English to be as much alike as possible. In fact we want them to be exactly alike, except for independently necessary parameters. So in fact if there's a delete-pronoun rule in Italian, as there apparently is, we want it to be universal. Let's assume that that's right and let's assume the universal rule Delete Pronoun. If it doesn't work in other languages let's suppose that there is some other property, already established for other reasons that will prevent it from working. In fact this rule should be part of the general "Avoid-Pronoun" principle I have been waving my hand at, which, if it was properly formulated, could have some consequences and the delete-pronoun rule as a special case. Recall that we have a structure like (60):

(60) NP INFL VP

As everybody knows there's some kind of relationship between the relation between INFL and subject and the relation between clitics to objects. In particular because the INFL is really PRO, that is AG which is just a collection of features, person, gender, number, index, case. Now that's essentially what clitics are, person, gender, case, number. So Clitics and INFL have to be the same element, that is they both are PRO, more precisely indexed PRO's. Now let's look at verbs. Thinking about verbs rather abstractly it looks like the following: Verbs assign case through subcategorization, they have subcategorisation features, $F_1, F_2, \dots F_n$. Let's assume then that it's the subcategorisation features that assign case, not the verbs which only appear to assign case because they have these features and the verb inherits this ability to assign case from them. So in fact let's assume that subcategorisation features have case, it's part of their meaning. Sometimes they have case inherently. Sometimes they may just gain a case by the fact that a structural case is assigned in the position that they control and then it goes on the subcategorisation feature. This is sort of a notation, but I want it to turn out that the subcategorisation features have the case, inherently (like "I assign inherent case") or via control (like "I happen to control that NP which gets objective case, for structural reasons, therefore I have case"). I am not going to worry about the proper formulation, it's all transparent.

Next point: Plainly languages are going to differ according to whether they have clitics or not. Now let's assume that the clitic is just a spelling out of the strict subcategorisation feature, along with AG, which is just PRO. So we have a rule like (61):

(61) $\left(\begin{smallmatrix} V \\ F_1 \end{smallmatrix} \right) \Rightarrow (v \alpha v) / \dots \beta$

β is the environment which is governed by the strict subcategorisation

feature. I'll use Rouveret and Vergnaud's superscript notation for government. So if the strict subcategorisation feature governs a NP rule (61) can apply where α is the agreement features of β^t , where the the agreement features are person number, etc i.e PRO. So this is in a way CL-PL, except that it doesn't tell you whether it's a PRO or a pronoun. Of course the β^t is only defined if it's pronominal, although we don't have to stipulate that. Of course we have to add something for the other clitics en and y in French. You could extend the approach by saying that the features of the prepositional phrases or adverbial, whatever they are, get attached to V as though they were . Or you could make the movement story about them. Here comes a question of facts. The Romance languages apparently differ very much as to whether you actually have an empty position in this β position. So, for instance, apparently, Argentinian Spanish keeps the pronoun always and standard Spanish only keeps it in datives and French loses it all the time. So there's a lot of variation depending on whether β is actually realised. In these terms this is going to mean whether β is a pronoun or a PRO. If it's a PRO you don't see it. If it's a pronoun it's there. Now the question is whether you have the same variation with respect to y and en. Depending on the answer to that question, you want to add these one way or another. If it turns out that they are always empty, then the obvious thing to do is to say that you really have a movement rule from the proper position into α . If it's not always empty you can do it like in (61). En and y of course are very different, they are not strictly subcategorized. So let's keep to the pronominal ones.

I am assuming that clitics have indices, i.e that α can be indexed, just like AG can be indexed. I'll assume further that here there's no language specific variation: If a language has clitics it has indexed clitics, although if a language has AG, which is just PRO inside INFL, it may or may not have indexing. The indexing is free but of course if you pick the wrong index, the sentence will be out, as usual. This in fact means that PRO elements are treated like NP's, even when they are inside verbs, like clitics. Also, if you think about it, everything is symmetrical, that is the relation between NP and AG is the same as the relation between NP and clitic, which is something that we want. Note that if the clitic has the feature, the clitic governs and the verb no longer does, i.e in

(62) ($V^{CL,V}$)

the outer V governs but not the inside verb. In other words in (61) α is the strict subcategorisation feature to which the AG is added. So the clitic is just the strict subcategorisation feature which has absorbed by assimilation or something like that the pronominal features of the NP which it is strictly subcategorized, hence that had to be pronominal. Notice that this is simply a way of formulating CL-PL which still allows for the option that you don't move the pronoun, which seems to be the case in lots of languages.

(Richie notes at this point that the "doubling" seems to occur only when the pronoun left behind is contained in a PP, even Argentinian Spanish where the direct object is preceded by a preposition. Chomsky says that if it turned out that CL-PL is always movement the rule could be modified since you can allow the choice of pronoun or PRO and a general principle could perhaps tell you you're only allowed to have pronoun if you have a PP)

Recall next that I was taking the index of the AG to be the feature that distinguishes, say, French and Italian. But here that distinction cannot exist since clitics always have an index. Now let's get down to the principle of delete pronoun. A possible formulation for this would be

$$(63) \text{PRO}_i \Rightarrow (\text{NP}_i^e)$$

(63) applies everywhere. As formulated the rule says that PRO keeps its index but loses its case, in other words (63) is really something like

$$(64) (\text{NP}^\alpha) \Rightarrow (\text{NP}^e)$$

where $\alpha = \text{PRO features}$

We can also assume that if there was a phonetic matrix in NP the rule wouldn't apply because you never delete a phonetic matrix. Let's assume further that (64) is a language universal. It hadn't occurred to me but when Richie was talking the other day he pointed out this odd phenomenon about Icelandic I realised that the description allows for the option that you have a case-marked PRO as long as the Delete-pronoun rule applies to it. Remember that the theory says that PRO cannot be governed, i.e. case-marked at LF, but of course there's nothing to stop PRO from being governed and case-marked previously: You could have a base-generated PRO which is governed and gets case and then undergoes the delete-pronoun rule so that by the time it gets to LF it's not PRO anymore. From what Richie said about Icelandic you could conceive that what is happening is that you have PRO, it gets oblique somewhere and there's an adjective that agrees with it by some agreement rule then PRO undergoes (64) and the resulting structure is fine. There are a fair number of languages where you actually get an oblique subject in indirect questions. So you get things like

(65) It is unclear what PRO to do

where you can get an overt subject instead of PRO with an odd case. Russian gets something like this ^{with dative} and Hebrew has a construction where you have something like "I don't know what is on me to do". Of course, under the present theory that's going to be needed: If you ever want to express the subject it's got to be case-marked somehow and since you can't get it from government then you'll have to apply some special case-assignment rule. So perhaps things like the Hebrew and Icelandic stuff aren't completely unrelated in having some odd way of keeping the subject in making up a case for a subject in an ungoverned position with the difference that in the case of Icelandic since you have PRO

it falls under (64). Of course this is probably wrong because our knowledge of Icelandic is so weak, but perhaps it's OK.

(Richie notes that the description has to be formulated so as to disallow control in nominative)

Now if we have a language with clitics and we have PRO in β then we drop the PRO by the PRO-drop rule in (64), then if the clitic has an index it'll properly govern the resulting empty NP (NP_i^e) and also bind it so everything is OK since (NP_i^e) meets the Binding theory, it's bound in its governing S and meets the Empty Category Principle since it's bound by the clitic. So if a language has clitics you can delete the PRO. Notice that I am trying to push to full generality. That is if a language has a place where you can put pronouns in then you can put PRO in the same position. Of course in a language like English you'll never have PRO because there's nothing ever to control it. But in a language like French you can sometimes have PRO since you have clitics. In Italian you're going to have it much more freely because not only it has clitics on verbs but also in effect it has clitic on the inflection (i.e AG with an index). So if a language has clitics then the PRO element can drop i.e the rule in (64) can apply. In English you can't see it happen. Comparable remarks hold for the relation between AG and NP: If a language has an indexed PRO in its INFL, then once you delete the subject by the Delete Pronoun rule you get

$$(66) (NP_i^e) \dots AG_i$$

and NP_i^e will be governed and also be bound and empty elements have to be bound by the binding theory in their governing category.

(Richie notes that with respect to disjoint reference it's the PRO, not the clitic that enters into it, so you don't want it to be gone when the binding conditions relevant for disjoint reference apply. Chomsky notes that if that's a real property then the description has to be changed.

Jean-Claude Milner notes that the analysis could be extended to subject clitics if you assumed that they are the spelling out of AG: If so the empty subject would be governed and bound. Chomsky agrees and notes that this would be symmetrical to the object case.)

Now assuming that the above is correct, what can we say about the case of the missing subject? I would like to say that the underlying form in Italian is

$$(67) PRO V \dots$$

i.e you pick the option of PRO instead of pronoun, which you're always allowed to do in any language; Then by the Pre-Drop rule (67) is turned into

$$(68) (NP_e) \dots V \dots$$

Then since the empty NP is properly bound and governed it's OK. If you tried to do it in French you couldn't do it, except perhaps, as Milner pointed out

A summary of Chomsky's Pisa lectures

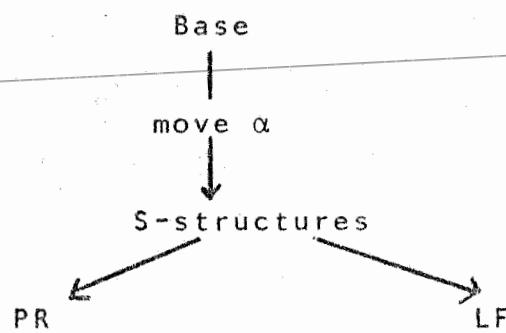
[Introductory note: this is an unauthorized account of Chomsky's Pisa lectures (Pisa, April 24-26, 1979). Critical notes and all errors are my responsibility, and I am indebted to Hans Bennis, Allied Blom, Saskia Daalder, Frederiek van der Leek, David Lightfoot and Henk Verkuyl for the informal discussions that contributed much to my understanding of my own notes. The reader should be warned that the following account possibly gives a somewhat distorted picture. Chomsky's own worked out version (Lectures on Core Grammar) will probably be published by Foris Publications this fall. - Jan Koster].

Lecture 1 (April 24)

1. Theoretical background

Core grammar: base, move α. The number of core grammars may be finite, in which case there is no theory of learnability! The following model is presupposed:

(1)



(base-generated structures are called D-structures, outputs of move α are called S-structures. Recently S-structures have also been called surface structures, but the latter term is confusing because S-structures contain traces, in contrast with traditional surface structures; PR = Phonetic Representation; LF = Logical Form).

For the base rules, a version of \overline{X} -theory is assumed: $\overline{X} \rightarrow \dots X \dots$, where $X \in \{\mathbb{N}, \mathbb{V}\}$. The following base rules are assumed:

- (2) (i) $\bar{S} \rightarrow \text{COMP } S$ (iii) $\text{INFL} \rightarrow \left(\begin{smallmatrix} +\text{Tns} \\ -\text{Tns} \end{smallmatrix} \right) \text{ AG}$
(ii) $S \rightarrow \text{NP } \text{INFL } \text{VP}$

New in these rules are INFL (for Inflection) and AG (for Agreement). AG includes features for person, number, gender, and case, and will play a prominent role in the third lecture, where an index of AG will govern the (phonologically) null subject in languages like Italian.

The structure of COMP is supposed to be as follows:

In many languages, a Wh-phrase can be followed by a complementizer: cf. ME: which that, etc. see F & C, OB. For the optionality of elements in COMP, there are three theories:

- (4) (a) F&C: free deletion in COMP + filters
 (b) OB : delete wh-phrases as much as possible
 up to recoverability
 (c) Pisa: optional expansion of COMP-extension rules

The third option (c) is adopted for these lectures. [The motivation for this choice wasn't quite clear. Symptomatic for a tendency to get rid of deletions?]. Traces in COMP don't have case and can always be deleted (special case of recoverability: there is nothing to recover).

2. Government

[The notion of 'government' is perhaps the central notion of the Pisa-theory. For two categories, α and β , it can be defined as follows:]

- (5) α governs β iff α minimally c-commands β
 and there is no \overline{S} or NP between α and β

$\alpha = [+N, +V]$, or $+Tns$

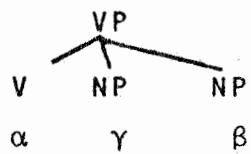
The notion of c-command is familiar from Reinhart 1976 (A node α c-commands a node β if neither dominates the other and if the first branching node dominating α dominates β). The addition of the word "minimally" creates some problems and leaves us with some unanswered questions. The notion of "government" falls more or less together with the notion of strict subcategorization, and the purpose of

definition (5) is clear: a lexical category (or Tns) governs its sisters; it also governs the non-sisters that it c-commands if there is no intermediate \bar{S} - or NP-boundary, and if there is no other lexical category or Tns that governs the non-sisters in question. The question is to what extent the intention is realized by the following definition:

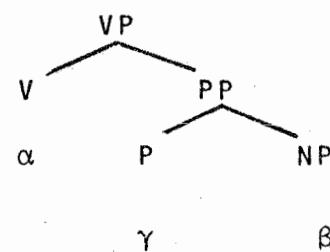
- (6) α minimally c-commands β iff (i) α c-commands β
(ii) there is no γ such that
 α c-commands γ and γ c-commands
 β and it is not the case that
 γ c-commands α

The definition can best be illustrated with some examples (assuming that we understand the definition):

(7)



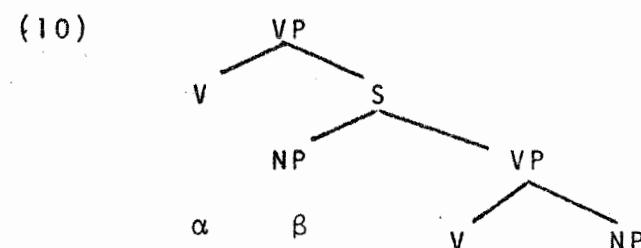
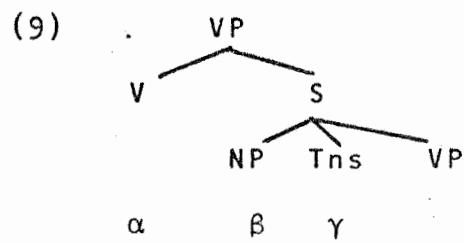
(8)



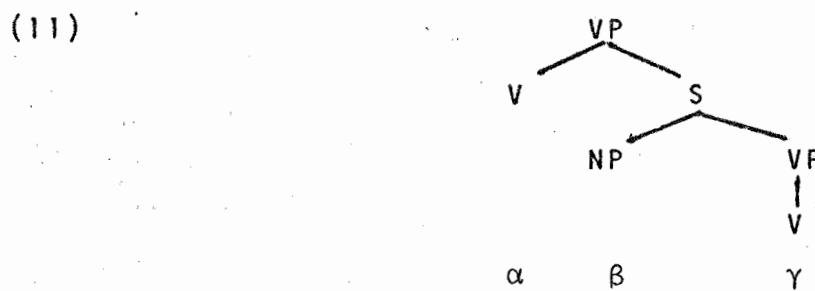
In (7), the V governs both NPs; they are c-commanded by the V, and even minimally c-commanded by V as required: the NPs don't qualify as possible γ 's because both c-command α ($=V$), which is forbidden by the last part of (6ii). In (8), the V does not govern the right-most NP. In this case, the intermediate P qualifies as a γ , because α ($=V$) c-commands P, and P c-commands the NP (β), and it is not the case that P c-commands V. In other words, V does not minimally c-command the NP, because it is in the domain of another governor, namely P.

Of special interest is the fact that a lexical category can govern a category (as far as we see only the subject) of its sentential complement iff there is no \bar{S} , but an S-boundary (see the stipulation with respect to \bar{S} (and NP in (5)). It is not quite clear how to account for the difference between \bar{S} and S. Chomsky suggested \bar{S} -deletion for the exceptional governing (and case-marking) into an embedded clause (perhaps for the case that S is not L-contained by \bar{S}) [for L-containment, see Conditions on Transformations (108)]. It is also possible that the difference is a matter of subcategorization.

zation, i.e. certain items might select \bar{S} , and others S. In any case, it is not unnatural to "neglect" \bar{S} if there is nothing in COMP so that \bar{S} falls together with S. According to the def. (5), it is possible to govern the subject of an S (and not of an \bar{S}). Consider, for instance:



In (9), the V cannot govern the NP (subject of the embedded S) because this NP is governed by Tns (Tns qualifies as a γ in the sense of (6ii)). In (10), there is no Tns in the embedded clause, so that the V of the matrix governs the NP subject of the embedded S. If there is an \bar{S} complement instead of an S, government is never possible according to (5). So far things work out smoothly with the definition of government in terms of minimal c-command. There are some problems however. Suppose that the V in the embedded clause of (10) is intransitive:



If (6) is taken literally, the matrix V does not govern the NP of the embedded clause because the V of the embedded S is a γ in the sense of (6ii), as observed by Frederiek van der Leek: the embedded V is c-commanded by the matrix V, and VP is not a branching node so that the embedded V c-commands the NP; and, finally, the embedded V does not c-command the matrix V. This complication is due to the definition of c-command: the non-branching character of the embedded VP is responsible for the fact that the lowest V c-commands the NP. The problem can easily be solved by changing the definition of c-command in the relevant sense.

Another problem arises in headless relatives in German, as discussed by Groos and Van Riemsdijk. In these cases, case assignment affects the wh-phrase in the COMP of the relative. So, this wh-phrase must be governed by the matrix verb:

- (12) Ich habe gesehen wen er gesehen hat

If the wh-phrase is governed by the matrix verb, we have the following situation:

This type of government is not allowed by (5). One might say that the leftmost (topmost) NP is governed by V, and that case "percolates" down to the wh-phrase. According to Chomsky this would be analogous to the fact that case "percolates" down from a governed NP to its head. But the situation is not fully analogous. But how serious is this problem? . It is reasonable to say that the wh-phrase (wen in (12) for example) gets its case in the embedded clause ("inherited" from the object position of the rightmost occurrence of gesehen; for "inheritance" of case, see the section on case-assignment below). In a sentence like (12) it does not seem necessary to say that wen also gets case from the matrix V. We only have what is usually called a matching effect. What Chomsky calls "percolation" is a possible description of the matching effect. Matching does not necessarily imply government. Thus, (13) may be changed in the sense that government works just in the normal way, while matching (or percolation) is taken as an independent issue that has nothing to do with government:

- | | | | | | | | |
|------|---|------------|--------------------|----------|-------------|---|---|
| (14) | v | [NP | [NP ^e] | [S̄ | [wh-phrase] | S |] |
| | | ↑ | ↑↑ | | | ↑ | |
| | | government | | matching | | | |

Apart from the notion of government itself, the following definitions form the core of the theory of government:

- (15) α is a governing category for $\beta =_{df}$ there is a γ
such that γ governs β and α contains γ

(16) α is a minimal governing category for $\beta =_{df}$ α is a
governing category for β which contains no other governing
category for β

(17) \nearrow Governing categories are NP and S (perhaps \bar{S})
Infinite
 $(\alpha = NP, S)$

The definitions (15) and (16) are straightforward. The selection of governing categories forms an important empirical issue, to which we will return.

3. Case-assignment

Case is assigned at S-structure by the following rules:

1. NP → nominative if the governing category = Tns (cf. 3rd lecture)
 2. NP → objective if " = [-N]
 3. NP → inherently case-marked as determined by idiosyncratic properties of [-N]
 4. Special rules like of-insertion in English (the destruction of the city)
 5. [NP NP - \bar{X}] → [NP NP genitive - \bar{X}]
 6. Condition: lexical α inherits its case from its trace

The categories V and P are [-N]. Thus, NPs governed by a P or V are assigned objective case. [It is not quite clear how these abstract cases are related to the "surface" cases of actual languages. For German, we might say that objective is realized as accusative, while genitive and dative objects are a matter of inherent case-marking]. Participles don't assign case:

- | | | |
|------|------------------------|--------------------|
| (18) | [[+V] en [V kill]] | NP
↑
no case |
|------|------------------------|--------------------|

Double objective constructions ([V NP NP]) are somewhat problematic. For English it can be assumed that the second NP has inherent case, and that the first has objective case if the verb is in the

active mood, and no case at all in the passive. An NP that is preposed by NP movement (passive) gets the case of its derived position (recall that case is assigned at S-structure). If an inherently case-marked NP is preposed, the result is a case-conflict because a lexicalized category inherits case from its trace. This explains why (19) is out (at least for many speakers):

- (19) *The book was given the boy t

The book is assigned nominative case because it is governed by Tns. But it also inherits the case from its trace, which entails a case-conflict. In a language like German, only accusatives are passivized. This can be explained by assuming that accusative is the only structurally assigned case (objective) and that the other cases for objects are inherent. It should be noted that "passive" is not a unitary phenomenon. Certain languages, like English, have NP movement, while others (cf. Arabic, in a paper by John McCarthy) have morphological passives [see also Jan Koster, Locality Principle in Syntax, 3.2.2.2.]. Cf. also the difference between John was taught French and John was untaught... Kayne has proposed to assign inherent case in the base. Does this make a difference?

To guarantee a proper application of the case-assignment rules we need the following case-filter:

- (20) *[_{NP} -case phonetic matrix] CASE FILTER

This filter does not apply on the left of the grammar (mapping to PR) but it is closer connected to LF. In fact, it can be assumed that it applies immediately after case-assignment. This is [?] to guarantee that all variables are properly case-marked:

- (21) [_{NP}e] is a variable iff it is case marked EXTENDED
CASE FILTER

A variable (vbl.) is defined as follows:

- (22) vbl. =_{df.} [_{NP}e] which is bound (not in the logical
sense!) by an appropriate quantifier

[a category is bound if it is coindexed with a c-commanding category].

It is crucial in the present theory that only traces of wh-movement are interpreted. The Extended Case Filter (21) entails that traces of NP-movement are never interpreted as variables. Passives involve participles which don't assign case to their objects (see above). Raising (seem, etc.) leaves a trace at the subject position of an embedded infinitival clause. Non-lexical subjects of infinitives (either trace or PRO) never have case, and are never interpreted as variables. Other non-variables ("traces without case") are the traces of wh-phrases in COMP. If it is assumed that wanna-contraction is blocked by variables but not by non-variables like PRO, we can get rid of Equi (self-deletion, see Conditions on Rules of Grammar). [Note that another deletion goes out of the window!]. [The relevant examples are as follows:

- (23)a. Who do you want [PRO to see t]
b. Who do you wanna see t
(24)a. Who do you want [t to see Bill]
b. *Who do you wanna see Bill

In (23), contraction is possible: neither PRO nor the (omitted) trace in COMP is a variable. In (24), the trace has case and is interpreted as a variable, which blocks case marking (cf. On Wh-movement and Lightfoot's paper)].

[The theory of variables presented here is the most remarkable feature of the Pisa lectures. It deviates from a certain logical tradition, in which ordinary personal pronouns can also be treated as bound variables (Quine, Geach, Montague, Partee, and many others, cf. also the last part of Chomsky's Conditions on Rules of Grammar, where an NP trace is still created as a variable (example (71)), and also a personal pronoun (examples (82) and (83))). Thus, in the logical tradition just mentioned and also in Chomsky's papers before On Wh-movement [?], there are at least the following kinds of variables (I am not sure about Chomsky's earlier positions for all cases):

- (25) Wh-traces
a. Who did you see t
b. For which x, you saw x
(26) Quantified NP positions
a. Everyone saw Bill
b. For every x (x a person) x saw Bill

(27) Focus

- a. BILL likes John
- b. the x such that (x likes John) is Bill

(28) Trace of NP-movement

- a. John seems [t to be a nice fellow]
- b. John, x seems [x to be a nice fellow] Conditions on R. (71)

(29) Personal pronouns (antecedent:quantified NP)

- a. Every soldier has his orders
- b. For all x , x a soldier, x has x 's orderd Cond. on R. (88)

(30) PRO (antecedent:quantified NP) (no examples in Cond. on R.)

- a. Everyone tries [PRO to go]
- b. For every x , x a person, x tries [x to go]

In the Pisa lectures, there are only variables in the sense of (25-27). NP-movement traces (28), personal pronouns (28) and PRO (29) are never interpreted as variables. They can be coindexed with a variable (bound in the non-logical sense), but they cannot be bound variables themselves. The distinction between variables and non-variables along these lines seems to be motivated by the following fundamental thesis (often repeated in Pisa):

(31) bound variables behave like names with respect to anaphora

Thus, in the Pisa theory (30a) is not interpreted as (30b) but as:

(32) For every x , x a person, x_i tries [PRO_i to go]

As will be shown later, PRO and personal pronouns do not behave like names (or like case-marked traces). For this reason, they are not treated as variables. In spite of this, it is not quite clear why he cannot be interpreted as a bound variable if it is linked to everyone in:

(33) Everyone thinks that he is happy

It cannot be interpreted as a referring expression (like bound variables in general) if everyone is the antecedent, and the binding relation is dependent on c-command just as the binding relation between a wh-phrase and its (case-marked) trace. So, under an alternative view PROs and personal pronouns can be interpreted as bound

variables if they are c-commanded by a quantified expression. In this sense, c-command is the grammatical analogue of the logical notion of scope. Chomsky's view is based on conceptual grounds (Variables are unspecified names) and the fact that variables in the sense of (25-27) but not in the sense of (28-30) behave like names. I think Chomsky's view is very interesting and coherent. But at the other hand it seems a little strange that an NP that is coindexed with a bound variable does not become a bound variable itself in LF (in spite of the fact that it is a non-referring expression). Personally, I feel that the whole issue needs considerable clarification].

4. Theory of indexing

There are two coindexing procedures:

- (34) (1) coindexing with movement
 (2) index freely

The first coindexing is familiar from On Wh-movement and subsequent papers. In Raising constructions, for instance, the moved NP_i leaves an [NP_i e] with the same index (trace):

- (35)a. D-structure: NP seems [John_i to go]
 b. S-structure: John_i seems [[NP_i e] to go]

In all other cases, indices are assigned freely. So, it is just an accident whether two categories end up with the same index or not. The results of free indexing are "filtered" by an independent theory of binding (see next section). [I suppose that (1) and (2) are ordered; if not, the NP subject of the matrix already has an index. The ordering follows from the organization of grammar: movement applies at the top part, free coindexing on the right].

This theory is a step away from Lasnik's theory (Remarks on Co-reference in LA). Consider the following sentence:

- (36) John says that he is happy

Lasnik only has a non-coreference condition. This condition does not apply, so that the interpretation of he is free in (36). He can refer to any person (male, singular), thus also to John.

In Chomsky's theory, he has either the same index as John, or a

different index. So, coreference is not only allowed, but it can be explicitly indicated by coindexing. If a pronoun is coindexed with something it is called proximate, obviate otherwise. The terms proximate and obviate are taken as descriptive terms here (in certain languages the distinction requires pronouns with different phonetic shape).

In general, we have the following kinds of NPs:

- (37) NPs

 - (1) anaphors: NPs that are lexically identified as
anaphors:reflexives, reciprocals, etc.
 - (2) pronominals (a) pronouns:have a phonetic matrix
+ features (person, gender, etc.)
(b) PROs:pronouns without phonetic
matrix but with features
 - (3) other lexical NPs (names, etc.)
 - (4) traces

PRO is proximate if controlled, obviate otherwise (= arb in OB).

5. Theory of binding (LF)

This is the heart of the present theory. It consists of three parts:

- (38) A: If α is an anaphor or lacks a phonetic matrix, then:

 - (1) α is a variable
 - or (2) bound in every governing category

B: If NP is case-marked, then:

→ (1) anaphor
or (2) free in every governing category

C: If α is pronominal, then free in every minimal governing category

[although the application of this theory is straightforward in many cases, the way it is formulated is not quite perspicuous. Why is it not simply formulated as follows:

- (39) A: bound in every governing category: anaphors, trace (-case)
B: free in every governing category: case marked NPs
C: free in its minimal NP or S:pronominals

This formulation clarifies the classification (of NP-types in terms of binding) that is implicit in (38). Is essential information from (38) missing in (39)? In any case, the classification seems rather natural: C is for pronominals, B for names (and other non-anaphoric lexical NPs) and variables (see 31!), and A for bound anaphors and traces of NP movement (and Wh-trace in COMP). A seems less natural. What have anaphors and traces of NP movement in common?

In 39C governing category has been replaced by S or NP, because PRO does not have a governing category; 38C does not seem quite correct in this respect.]

It follows from (38) that PRO is not governed, and that Trace of NP can't be case-marked:

- (40)a. -assume that PRO is governed
-then there is a minimal category α
in which PRO is governed
-PRO must be bound in α (39A)
-PRO must be free in α (39C)

The last two lines form a contradiction, therefore:
PRO cannot be governed

- b.

 - assume that NP-trace is case-marked
 - then either anaphor or free in every governing cat. (38B)
 - but trace always coindexed (not free) in gov. cat.
 - thus anaphor
 - but it is not an anaphor (by definition)
 - therefore, it cannot be case-marked

Ne falls under { 38 A because it is null phonetic matrix
38 B because it is case marked

A- must be found in every g-cat. (can't be a vbf)

B - must be full in every g - eat (can't be 'empty')
better - 'normal'

Lecture 11 (April 25)

[This lecture gives illustrative examples for the theory discussed in the first lecture. It is, by the way, clear why (39) does not express the same as (38). (38) is not only a classification (like 39) but also the basis for certain deductions. It follows, for instance, from (38) (but not from 39) that PRO can't be governed (see 40a). Certain implications of (38) can be clarified, I suppose, when it is assumed that categories can be trivially free or bound in a governing category, namely in the case when there is no governing category. A second assumption is that the CASE FILTER (20) not only refers to NPs with a phonetic matrix but also to variables (this was in fact added to (20) during the first lecture). So, (20) should be read as follows:

$$(41) \quad *[\underset{\text{NP-case}}{S} \text{ phonetic matrix (or variable)}] \quad \text{CASE FILTER}$$

With these two assumptions, (38) properly applies when PRO is the subject of an infinitival clause, or when a trace is in COMP. Consider for example:

$$(42) \quad \text{It is impossible } [\underset{S}{\text{PRO}} \text{ to go home}]$$

If PRO is α in the sense of (38), we have the following situation. Condition (38C) is trivially fulfilled: PRO is pronominal and it is (trivially) free in every minimal governing category because there is no governing category for PRO in (42) since it is not governed. Condition (38B) does not apply because PRO (in 42) is not case-marked. Condition (38A) does apply (because PRO lacks a phonetic matrix). Suppose that PRO (in 42) is a variable. Then the sentence is filtered out by the CASE FILTER (41), because the ungoverned PRO in (42) does not have case. So, only the second possibility of (38A) applies. This condition is trivially fulfilled: PRO (in 42) is bound in every governing category because there is no governing category. A trace in COMP gets a similar treatment:

$$(43) \quad \underline{\text{Who}}; \text{ do you think } [\underset{S}{\text{[NP}_i \text{ e]} \text{ [Bill saw [NP}_i \text{ e]]}}]$$

Again, only (38A) applies (the trace in COMP is neither case marked

nor pronominal). If the trace in COMP is a variable the sentence is ruled out by the CASE FILTER (41) (the trace in COMP does not have case). The second condition (38A (2)) is again trivially fulfilled because the trace in COMP is not governed and lacks therefore a governing category].

We can derive the following theorems:

- (44)a. PRO is ungoverned (cf. 40a)
 b. [N_pe] is a variable iff case-marked

The second condition involves a biconditional and has therefore two parts ($L \rightarrow R$, from left to right; $R \rightarrow L$, from right to left):

- (45)a. L→R: if $[_{NP}e]$ is a variable then it is case-marked
 b. R→L: if $[_{NP}e]$ is case-marked then it is a variable

(45a) follows from the CASE FILTER (41); (45b) follows from the binding theory (38).

The following cases will be illustrated:

- (a) PRO (in S and NP)
 - (b) Pronouns (in S and NP)
 - (c) Case-marked non-anaphors (lexical items + variables)
 - (d) Trace of NP

(a1) PRO in S

PRO is ungoverned at LF. The binding theory entails that PRO is only allowed in the subject position of infinitives:

If the matrix head allows \bar{S} deletion, the subject position of infinitives is governed. This explains the following contrasts:

- (47) a. *It is certain [_S PRO to go home]
 b. John_j is certain [_S [_{NP}_j e] to go home]
 c. It is impossible [_S [PRO to go home]]
 d. *John_k is impossible [_S [_S [_{NP}_k e] to go home]]

In (47a), PRO is governed because certain triggers S̄ deletion. Therefore the sentence is out. Impossible does not delete the S̄. As a

consequence, PRO is allowed: the \bar{S} boundary blocks government of the embedded subject position. Note that (47d) is not ruled out by the binding theory but by RES(NIC) (residue of the NIC, to be discussed below, which says that $[NP_e]$ has to be governed).

[I can't help being sceptical about this explanation and about similar use of the $S-\bar{S}$ alternation in Bresnan's thesis (chapter 3 ?) The $S-\bar{S}$ distinction is rather crucial in the present framework and is part of a strategy to maximize the distinction between TRACE and PRO. Thus, we have John seems [$_S$ TRACE to go] and John tries [$_{\bar{S}}[S$ PRO to go]]. TRACE and PRO differ here in three respects: (i) TRACE is governed, PRO isn't (ii) TRACE requires \bar{S} deletion, PRO does not, (iii) PRO differs from TRACE in that the former (but not the latter) has features (cf. 37, 2b). These three differences are almost completely unmotivated. If there was no difference between TRACE and PRO in infinitives, we could simplify the definition of 'government' (recall that the difficulties with minimal c-command were caused by the fact that government sometimes goes down into embedded clauses). The TRACE-PRO distinction is also an argument for having two indexing procedures: movement and free indexing. In other words, the TRACE-PRO distinction complicates the grammar in more ways than one. Since the distinction in question has little motivation to begin with, it seems better to get rid of it (unless more evidence can be found). This is, of course, a central issue and I hope to return to it elsewhere. To mention just a few problems: TRACE is governed by a verb (seem), but it is not assigned case. To block application of the second rule under 3. (p.6 above), one has to have recourse to the notion of "strict subcategorization" (seem does not have an object). But since "strict subcategorization" is almost the same as "government", one notion seems enough (the objects of a V are the NPs governed by that V). A problem with two indexing procedures (movement and free indexing) is that one cannot avoid double derivations in a non-ad hoc fashion. Consider, for instance, John; was killed [NP_e]. This sentence can be derived in two ways: (i) by movement, (ii) by non-expansion of the object of kill and free indexing. Neither the TRACE-PRO distinction nor the need for two indexing procedures seems established (at present at least)].

(a2) PRO in NP

* John saw {^{PRO}_{NP-e}} John is expected {^{PRO}_{NP-e}} to
John tried {^{PRO}_{NP-e}} down

The following cases are impossible because PRO is not allowed in a governed position:

- (48) a. *I think that a picture of PRO ...
b. *The destruction (of) PRO ...

Possessives deserve some special attention. The following rules is assumed:

- (49) $NP \rightarrow [_{NP} NP + Poss] / -- \bar{X} (= \bar{N} \text{ or } \bar{V})$

The special element Poss is the governor in this case. Thus, the following case is ruled out:

- (50) *John likes [_{NP} PRO's book]

This sentence is ungrammatical because PRO is governed by Poss ('s). If $\bar{X} = \bar{V}$ (in 49) we do have the possibility of PRO:

- (51) John prefers {^{my}_{PRO}} writing the book
 {_{his}}

This difference is accounted for by making the case-assignment by Poss optional in (51) but not in (50). In general, there seems to be something like an avoid pronoun principle: PRO and pronoun (often) in complementary distribution. If PRO is possible, a (proximate) pronoun must be avoided. That is the explanation for the alleged ungrammaticality of John; prefers his; writing the book. [According to all available native speakers of English, this sentence is grammatical, however]. Apparently, the avoid pronoun principle is not without problems. (Cf. also: John bought a book (for) him (DR?), (PRO) each other (*) to read. A complication here is that this sentence is structurally ambiguous (cf. G. Horn in LA).

(b) Pronouns (in S and NP)

Most examples are straightforward (G = Governing category):

- (52) [_G John saw him]

Here we have DR (Disjoint Reference) between John and him. Suppose they had the same index. Then him would not be free in G, which is required by (38c). The following sentence has a reading with John

and he coreferential:

- (53) John_i believes that [_G he_i is sick]

In this case, the pronoun he is free in its minimal G (=S). Compare:

- (54) [_G John believed [_S him to be sick]]

Here we have DR between John and him because him is only governed in the matrix S (=G), in which it is not free (as required by 38C).

Pronouns in NPs can be coreferential with antecedents outside of these NPs:

- (55)a. John_i saw [_G my picture of him_i]
b. John_j saw [_G his_j picture of me]

In both cases, the pronoun is free in its minimal G (=NP).

Special problems are posed by sentences like the following:

- (56) John_k kept it [_{PP} near him_k]

This problem can perhaps be solved by extending the notion "governing category" to the PPs in question.

(c1) Case-marked non-anaphors (lexical items)

Case-marked lexical NPs have to be free in every governing category G

- (57) *[_G John_i saw John_i]
*[_G John_i said that [_G John_i]]
*[_G John_i said that [_G Mary thought that [_G John_i ...]]]

etc.

(c2) Case-marked non-anaphors (variables)

The binding theory guarantees COMP-to-COMP movement. Suppose that an NP moves from COMP to a non-COMP position (as discussed by Robert May):

- (58) [_{COMP} [_{S₂} ... vbl. ... [_{COMP} [_{S₁} ... vbl. ...]]]]



This derivation is ruled out by (38B, 2): a case-marked NP must be free in every G. In (58), the vbl. (case-marked) in the most

deeply embedded S is free in this S₁, but not in S₂. [an NP is free if it is not bound; an NP is (argument)bound iff it is an argument coindexed with a c-commanding argument. Arguments are base-generated NP-positions. Thus, (38) presupposes argument binding. Note that the COMP-to-COMP condition for prepositional Wh-phrases requires some further assumptions. Apparently, the binding theory (38) abstracts away from Pied Piping effects. Other examples involving variables are:

- (59)a. *John_i hit [NP_i e]
b. *Who_j e_j hit e_j

In both these cases the object of hit is case-marked. It can't be an anaphor, so it must be free in every governing category (38B), which it is not].

The following cases were discussed by Higginbotham:

- (60)a. *Everyone_i saw him_i
b. *He_j saw everyone_j

According to Higginbotham, these facts are explained by the contradictory index on the object (pronounce: "i-set-i"):

- (61)a. (x) (x_i saw him_{{i,{i}}}) (= 60a)
b. (x) (he_j saw x_{{j,{j}}}) (= 60b)

In Chomsky's view, there is no DR with variables. This would entail a difference between (60a) and (60b). In (60a), the pronominal him is not free in its minimal governing category (S), but he in (60b) is free. Compare:

- (62)a. Carter voted for everyone on the Democratic list
b. Someone in Plains, Georgia, voted for Carter

In (62a), everyone, etc. can include Carter (everyone is a quantified NP; the universal quantifier binds a variable here. No DR with variables. The name (Carter) is free in its governing S).

In (62b), the phrase including someone does not include Carter. Here we do have DR, because the name (Carter) has to be free in S. [Both the facts and the explanations seem questionable. Nobody was able to reconstruct the full issue from his notes. So, probably something is missing. It seems that in both (60a) and (60b) everyone

can include him (he). The binding theory does not predict a difference between (6)a) and (6)b). In (6)a) the pronominal him is not free in G, and in (6)b) the variable is not free in G. For both cases, the binding theory predicts ungrammaticality (38C and 38B, respectively). All in all, it has not become clear how variables can be made immune for DR].

Note that there is no strong cross over condition:

- (63) *Who_i [did he_i say Mary kissed e_i]

The ungrammaticality of such sentences follows from the binding theory: the rightmost e_i is not free in every G (38B), because it is bound by he_i.

The weak cross over cases do not follow from the binding theory:

- (64) ?*Who_j did the girl he_j saw kiss e_j

In this case, he_j does not bind e_j, because the former does not c-command the latter. [For the difference between strong and weak cross over, see Wasow's dissertation].

In general, the SSC and the NIC (with an exception, see next lecture) follow from the binding theory.

d. Trace of NP

Consider the following examples:

- (65)a. [G John was killed t]
b. [G John was believed [S t to be a fool]]

In (65a), John can bind the trace (t). The trace does not have case, although it is governed by kill (participles don't assign case). The trace is bound in G, so it can't be a variable and (38A, 2) applies. In (65b), everything is similar if we assume that the matrix S is the governing category (t is governed by believe in G). The following sentences are ruled out:

- (66)a. *[John was believed [G t was killed]]
b. *[John was believed [G NP to have killed t]]

In (66a), the t is governed by Tense. So, it is not bound in G (38A, 2). In (66b), t is a non-anaphoric case-marked NP. So, it cannot be bound by the NP in G, or by John in the matrix. Since there is no proper

operator phrase at all in this case, the sentence is ungrammatical.

The next case involves Raising:

- (67) [G John seems [S t to go]]

Seem deletes S so that the trace is governed by seem in the matrix (G). Again, we have a non-pronominal non-variable, which is properly bound in the only governing category (38A, 2). The following sentence is acceptable according to the binding theory, but bad for independent reasons:

- (68) ?John expected [G me to seem [t to go home]]

[Anaphors were hardly discussed during this lecture. Thus, Chomsky didn't return to problems like: they saw [G each others pictures]. This sentence is ruled out by (38A, 2), because the anaphor each other is not bound in G (= NP). At the GLOW lecture, this case was presented as an instance of markedness. The following problems were briefly discussed:]

- (69) ?They left with each other

It seems that reciprocals are better in object position. It is not clear why. The following is OK, however:

- (70) They are easy for each other to talk to

Certain cases of A-over-A follow from the binding theory:

- (71) *Whose did you see [NP t house]

Whose contains the governor (= Poss). So, the trace does not have case. Uncase-marked traces must be bound in every governing category, thus also in NP in (71).

• Lecture III (April 26)

According to the binding theory, variables are free in every governing category and (therefore) free from any effect of the NIC or SSC. Thus, variables do not behave like anaphors, but like names with respect to binding. This is what we really want (there is a close connection: names are instantiations of variables, variables are unspecified names). With respect to the NIC, there are some problems

- (71) *Who[do you wonder [how well [_s t saw Bill]]]

The NIC correctly predicts this ungrammaticality, but there is no immediate explanation under the present theory [the t is free in every governing category as required by 38B, 2]. This problem will be discussed as the RES(NIC) (the residue of the NIC).

1. Four problems for the RES(NIC) (the OB-position)

- (a) It does not apply in Italian (in all other cases, the binding theory applies without exceptions in Italian. Only difference with English: no NIC for wh-traces.

(b) Focus

In sentences like: John said that BILL liked Mary, BILL has to be replaced by a vbl. in LF (see Conditions on Rules, and 27 above). Thus, NIC is too strong in this case

(c) Broad Scope quantifiers

In sentences like: John said that some of the talks would be too difficult, the quantifier some is said to have broad scope, i.e. May's QR adjoins the the quantifier phrase to the matrix S (For some x (= talks) John expected that x would be ...) Again we have a vbl. in the subject position of a tensed clause, and it is not bound within its minimal \bar{S} , as required by NIC. Again, the NIC is too strong.

(d) Problems discussed in paper by R. Kayne

2. What has the RES(NIC) to deal with?

- (a) missing subjects in Italian: Ho trovato il libro

(b) free stylistic inversion in Italian: t V NP

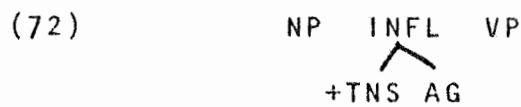

(c) long wh-movement in Italian (the violation of NIC in wh-movement?)

(d) resumptive pronoun strategy (discussed by Tarald Taraldsen)
Ecco la ragazza [S che mi domanda [chi -- possa]]
Here the girl that I ask who can

(e) *[that t]-filter does not hold in Italian

What these cases have in common is an empty subject position in a tensed clause, with no binding (wh)-phrase in COMP. So, why does the

NIC not apply here, as it does in English? The answer is that in Italian (and Spanish) but not in English (French, German or Dutch) the subject position can be bound by AG (= Agreement):



The assumption is that in Italian, there are indices on AG (but not in French, etc.). Together with a filter proposed by Pesetsky, and a reformulation of the notion of government, the hypothesis about AG indices can explain the difference between English and Italian. A preliminary formulation of RES(NIC) reads as follows:

- (73) RES(NIC): $\left[\begin{smallmatrix} NP_e \\ nom \end{smallmatrix} \right]$ can't be free (nom=nominative)

The Pesetsky filter blocks doubly filled COMPS:

- (74) Pesetsky filter: *[$\alpha \beta$]

A trace followed by that, for example, is forbidden by (74): *[t that]. In order to derive grammatical sentences, the filter has to be supplemented with an optional deletion rule that drops uncase-marked $[NP_e]$ at the level of S-structure (i.e. Shallow structure). Consider the following examples:

- (75)a. Who does John think [[t that][Bill saw t]]
b. Who does John think [[t] [Bill saw t]]
c. Who does John think [[t that][t saw Bill]]
d. Who does John think [[t] [t saw Bill]]

(75a) is filtered out by (74), unless one of the two elements in COMP is absent (t deletion, or optional expansion of -WH=that). (75b) is always OK (no that). In (75c), t cannot be deleted because of RES(NIC) (the trace in subject position has to be bound in \bar{S}). The sentence (75c) is ruled out by the Pesetsky filter (74). In (75d), the t cannot be deleted because of RES(NIC) (73).

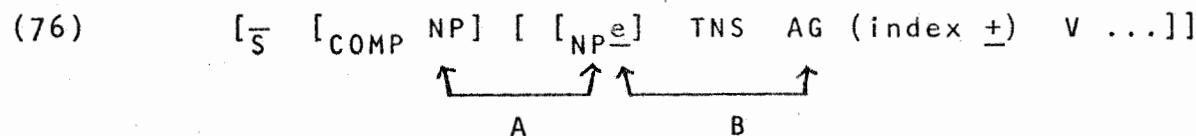
[we have omitted here some obscure cases discussed by Richie Kayne]. There are two problems with the that-trace filter in English (see F&C): (i) relative clauses: $[NP_e NP \text{ that } e \dots]$
(ii) clefts: It was the boy that e left

[As far as I see, these problems remain unsolved, also in what follows]

3. What has the RES(NIC) to deal with?

- (a) What does "free" mean in the definition of RES(NIC)?
- (b) Why is there an asymmetry between +nom and -nom?
- (c) Recoverability of deletion

The problems are solved by reformulating RES(NIC) in terms of a reformulated notion of government. This notion of government, which will be defined within a moment, is partly based on coindexing. There are two relevant forms of coindexing: from COMP to the subject position (of the tensed clause), and from AG to the subject position:



English has only coindexing from COMP to the empty subject (A). Italian also has indexing from AG to the subject (B). This explains the differences between the two languages.

The answers to the questions (a) and (b) are as follows. "Free" in the RES(NIC) means "governed":

(77) RES(NIC): df. $[_{\text{NP}} \underline{e}]$ must be governed
nom

The answer to question (b) is: there is no asymmetry. The notion nom can be dropped, because objects are always governed anyway; the option (+governed) only exists for subjects (depending on the coindexing possibilities to be discussed within a moment). So, we can drop nom in (77):

(78) RES(NIC): df. $[_{\text{NP}} \underline{e}]$ must be governed

What we need in fact for RES(NIC) is the notion of proper government:

(78) proper government

α properly governs β iff α governs β

if (i) $\alpha = [\pm N, \pm V]$

or (ii) α is coindexed with β

The def. of government guarantees that the coindexing is local

(minimal c-command, and no intervening NP or \bar{S}). The two possibilities are:

(79)a. [COMP_i] [S [NP_ie] ...]

b. [S [NP_ie] ... AG_j]

Since Italian has the second option (79b), the facts under 2. (p.21) are accounted for. If Italian lacks a binder in COMP the empty subject can still be governed by coindexing from AG. English misses this option.

It is no longer assumed that TNS assigns case (TNS is no longer a governor). This is somewhat problematic because we loose an account for the fact that subjects of -Tns clauses are not assigned case. [An even more serious problem is the following. The reformulation of government predicts that a sentence like John walks is ungrammatical because John is neither governed by a lexical category nor by coindexing with an element in COMP or AG. It is not so obvious how Tns can be missed as a governor. Another problem is that [NP_ie] in COMP is never governed. How do we block application of RES(NIC) in that case?]

The reformulation of government (proper government) entails that S is an absolute boundary for movement [at least in the case of NP-movement. Nothing was said during these lectures about Subjacency. Most island problems are left where they were]. Consider:

(80) John tried [\bar{S} [NP_ie] to leave]

The [NP_ie] in this case cannot be a trace, because it is not governed: the \bar{S} boundary precludes government by try. Since government for [NP_ie] is obligatory (RES(NIC)), a sentence like (80) cannot be derived by movement. We must have PRO in the embedded subject position.

[There were a few sketchy remarks about recoverability that we were not able to reconstruct. The lecture ended with some remarks on the nature of clitics].