Economy, Earliness and LF-Based Syntax

MICHAEL BRODY

Introduction

In this note I should like to show how some data that has recently received attention is quite straightforwardly and naturally explained if derivations start from LF, that is if S- and D-structure are derived from LF rather than the other way round. I shall not attempt at all to do justice to the analyses I draw on, but shall only discuss in a rather sketchy and programmatic fashion those aspects that are relevant for my purposes.

1 Economy
1.1 Main Verbs

According to Pollock (1989), the reason why in tensed clauses French main verbs raise to Inf but English Inf lowers to main verbs, has to do with the theta criterion. Inf is opaque in English: if a verb raises into it, theta role assignment by this verb is precluded. In French Inf is transparent; --raising a potential theta assigner to Inf does not prevent theta assignment, so no theta criterion violation results. Hence the derivation in (1a,b) and the ungrammaticality of (1c), --on the assumption that the adverb is positioned between Inf and the VP:

(1) (a) John often eats+Inf(opaque) spaghetti (Inf lowers)
(b) Jean mange+Inf(transparent) souvant des pommes (V raises)
(c) * John eats+Inf(opaque) often spaghetti (V raises)

This leaves (2) to account for, where Inf lowered in French results in ungrammaticality:

(2) * Jean souvant mange+Inf des pommes (Inf lowers)

Consider Chomsky's (1989) solution. He proposes that pre-S- structure V-raising, where it can occur, takes precedence over Inf-lowering because

shorter derivations are to be preferred over longer ones[1]. V-raising involves one step only while the trace of Inf-lowering has to be covered by LF through the later raising of the V+Inf unit. This round trip derivation[2] consists then of two steps and therefore is not chosen where V-raising is possible ie. --where Inf is transparent. Hence the ungrammaticality of (2).

This "least effort" condition is quite "intelligent" in the sense that in order to operate it has to be able to compare derivations and count steps in them. There are many technical problems and open questions involved in counting steps[3]. More importantly it implies global (counting steps) and trans-derivational (comparing the number of steps) inspection. Although this power is attributed only to UG and clearly would not apply in a language particular fashion, it still seems overly powerful. Furthermore, it is difficult to see how a reasonable theory of use and/or acquisition could be based on a competence theory incorporating transderivational principles.

An earlier interpretation of the least effort idea can be found in Chomsky (1986) who suggests that movement is a "last resort", it can apply only when forced by some condition. Thus for example NP-movement can only apply when forced by Case theory. The last resort condition is a more "primitive" least effort condition. Instead of counting and comparing steps in derivations, could we get the same result by a simpler last resort condition like the one in (3)?

(3) affect-alpha only if this is necessary to satisfy some condition

Suppose derivations start at LF. The LF of all the examples in (1) and (2) will be as (1b/c) with the V+Inf in the position of Inf, --otherwise the unbound, ungoverned trace of Inf is not covered at this level. In French between LF and S-structure nothing needs to happen since Inf is transparent. Hence by (3) nothing happens, (1b) is derived and (2) where both the verb and Inf lowered are impossible. In English Inf is opaque at S- structure[4] and this

---

[1]There are in fact two economy principles proposed by Chomsky: (a) that shorter derivations are to be preferred over longer ones and (b) that UG principles are applied whenever possible, language particular rules operating only where necessary to "save" derivations. Here I consider the first of these principles only.

[2]To use Pesetsky's (1989) felicitous term.

[3]For example, why does movement followed by deletion count as one step rather then two when Agr or infinitival Inf/Tns lowers and then deletes?

[4]I assume either that (a) opacity/transparency is checked for at S-structure only or, equivalently, that (b) the verb has to be in a transparent position throughout the derivation and English Inf is opaque only at S-structure, and becomes transparent at LF. As Pesetsky

will force the verb to lower between LF and S-structure excluding (1c). We can assume that Lasnik's filter will cause Inf to also lower and catch up with the verb. There are two possibilities here. If Inf lowers first and the verb is adjoined to it we derive the opaque configuration: [/Inf V+Inf]. If however the verb lowers first then we end up with the transparent structure: [/V V+Inf]. The latter option gives the correct S-structure representation. This accounts for (1a).

1.2 English Auxiliaries and Negation

Let us now turn again to the standard D-structure to S-structure to LF framework. Here negation appears to block LF raising but not pre S-structure raising:

(4)	(a)		John does+Inf not eat apples
	(b)	*	John not eats+Inf apples
(5)			Jean ne mange+Inf pas des pommes

In contrast to (1a) where Inf lowering is fine the S-structure in (4b) cannot be saved through V+Inf raising at LF, presumably due to the intervening negation. Arguably negation is a head, and of so LF raising in (4) will violate the HMC/ECP.

Why is then the French example with S-structure verb raising in (5) fine? Chomsky's solution in terms of gamma marking by and deletion of a head situated between the base position of the verb and Inf (his Agr-O) is highly problematic for a number of reasons, cf. Pesetsky (1989). The alternative, suggested by Pollock and adopted also by Pesetsky and others is to consider *ne* as the head of the negP in French corresponding to *not* in English. Since *ne* is a clitic, the verb can move into the *neg* head and the neg+verb unit can move higher into Inf without violating the HMC/ECP.

The problem with this solution as noted by Pesetsky is that auxiliary verbs in English can also raise past negation exactly like French main verbs:

---

points out under the alternative analysis where Inf is opaque both at S-structure and LF and only the adjunction mechanism is relied on to ensure transparency, the question arises as to why the round trip derivation cannot be completed before S-structure. Another problem in the round trip analysis is that if V+Inf raises and substitutes for Inf then Inf will be in an embedded position which, as Chomsky notes, raises the question of how it can satisfy the requirements of tense interpretation at LF. If on the other hand V+Inf raises to adjoin to Inf, then the verb will end up in an opaque position if Inf is opaque also at LF. Thus it seems that the assumptions (a) or (b) above are necessary also in Chomsky's analysis.

(6) (a) John has not eaten
(b) John was not seen

But *not* in English cannot in general be clitic, if only because we would then loose the explanation of (4b). Pesetsky is forced to stipulate that *not* is a head for the LF movement triggering the HMC/ECP in (4b), but a nonhead for the S-structure movement in (6).

Suppose that derivations start at LF and we have the least effort condition in (3). Nothing forces the auxiliaries in (6) to lower by S-structure, thus LF and S-structure will be identical in relevant respects by (3). Assuming that Lasnik's filter holds also at LF and thus Inf has to be supported at this level by a lexical head, the LF of (4b) must be (4a). Recall that the verb cannot support Inf because Inf is opaque. In (4a) nothing needs to happen hence the S-structure (4b) can not be derived from it.

2 Earliness

2.1 General

Pesetsky (1989) suggests an alternative to the economy principles of Chomsky (1989). He proposes that instead of preferring the derivation consisting of the smallest number of steps, we should choose the one where filters are satisfied earliest in the derivation. Abstracting from the details of what it means of satisfying a filter, in general this has the effect of favouring a derivation D in which rule R applied earlier to D', where it applied later. This would immediately give the observed order between V-raising on one hand and Inf-lowering followed V(+Inf) raising on the other, since in the latter V-raising applies later[5].

The earliness principle like the economy condition is undesirable since it involves global and transderivational power. The earliness principle seems to

[5]I have oversimplified Pesetsky's proposal here. He actually says that a filter with SD A=[a/1,...,a/n] is satisfied at level L for the purposes of the earliness principle iff for all a/i in A no element that is in a chain with a/i violates any grammatical principles at a level later than L. Thus Lasnik's filter is satisfied by V-raising at S-structure, but at LF only by Inf lowering. This is because although Inf itself is supported once it lowered, its trace, a chain-mate of an a/i (a/i = Inf) in the SD of Lasnik's filter violates the ECP at S-structure; and this is remedied only at LF. Hence V-raising should be preferred over Inf-lowering when both are possible.

Pesetsky attempts to bring the preference for UG rules over language particular ones (Chomsky's principle (b) in NOTE/1 above) under his earliness principle through postulating a language particular level of representation which he deems to be later than LF. I shall ignore this aspect of his theory here.

have the further drawback of being implausible for biological systems. It is easy to see why evolution would favour least effort type principles. A "lateness principle" would be compatible with the idea of getting a certain result with the minimum effort, the earliness principle clearly is not. Thus it might force an operation to take place which may become unnecessary at some later stage.

2.2 Wh-in-situ

Pesetsky (1987) observed that wh-phrases that ask questions whose range of possible answers is taken to be limited/given in the discourse or situational context of the utterance (D-linked wh- phrases) consistently differ syntactically from non D-linked ones. Thus for example a typical D-linked wh-phrase like *which man* shows no superiority/ECP effects when left in situ at S- structure:

(7) Mary asked which book which man read t
(compare: "*Mary asked what who read")

Since Huang (1982), it has been widely assumed that interrogative operators must move to an appropriate A'-position at LF. Pesetsky proposed that D-linked wh-phrases, thus eg. *which man* in (7), unlike non D-linked ones, need not end up in an operator position at LF. Thus they can stay in situ even at that level, leaving no trace subject to exclusion by the ECP. He noted that in a language like Polish that allows multiple wh-movement in multiple questions, a wh-phrase in situ must be interpreted as D-linked:

(8) Kto robi co
Who does what

By distinguishing D-linked and non D-linked wh-phrases in terms of their S-structure position, Polish corroborates Pesetsky's distinction. But also it raises the question of why a wh-in-situ in this language like for example *co* in (8) must be interpreted as D-linked? Pesetsky (1989) explains this in terms of his earliness principle. If *co* in (8) was not D-linked it would have to move to an operator position by LF. Since it is possible for it to move before S-structure (Polish has multiple syntactic wh-movement) it must do so by the earliness requirement.

Since we do not wish to adopt the earliness principle for conceptual reasons let us look for an alternative. Suppose that derivations start from LF and the "primitive" last resort condition in (3) is operative: move alpha only if some requirement of the grammar makes this necessary. Immediately non D-linked wh-in-situ is excluded in a Polish type language: non D- linked wh-phrases are in operator position at LF and nothing forces them to lower by

S-structure since Polish allows multiple wh-phrases in sentence peripheral position at S-structure.

3 The Wh-Criterion

As exemplified by (9) a +WH CP in English must contain a wh- phrase at S-structure:

(9) (a) John wondered who +WH you saw
(b) John wondered +WH you saw who

Rizzi (1991) suggests that the requirement that (non D-linked, interrogative) wh-phrase operators must be in a +WH CP also holds at S-structure in English. He formulates the wh-conditions as in (10) below:

(10) (a) a wh-operator must be in a spec-head configuration with an X/0 [+WH]
(b) an X/0[+WH] must be in a spec-head configuration with a wh-operator

If both a condition like (10a) and one like (10b) hold not only at LF but also at S-structure in English, then they can account for the 'residual' V/2 effects as for example in (11):

(11) Who did John see t

Assuming that the +WH feature is on Inf, the wh-phrase and Inf both have to raise to CP to be in the spec-head configuration required by (10) at S-structure[6].

---

[6]Rizzi argues that (10) at S-structure excludes also cases where an interrogative wh-phrase is in an inappropriate A'- position. Thus (10a) would rule out for example in (ia) where the wh-operator is in a -WH CP, or (iia) where it is VP-adjoined:

(i) (a) * I wonder +WH John thinks who -WH Mary saw t
(b) I wonder who +WH John thinks t -WH Mary saw t
(ii) (a) * Il a combien lu de livres
He has how many read of books
(b) combien a-t-il lu [t de livres]
How many did he read of books

In general if (10) holds at S-structure it predicts the effects of the generalization that wh-movement between S-structure and LF cannot start from an A'-position (cf. Aoun

However, in order to allow wh-in-situ Rizzi has to restrict (10) to wh-operators, that is to wh-phrases in A'-position[7]:

(12) wh-operator= wh-phrase in A'-position

As he notes (12) can hold only at D- and S-structure, at LF "it is superseded by a stronger principle according to which all elements endowed with intrinsic quantificational force are operators at this level, and must be moved to an appropriate scope position" (p.22). But if the notion of operator is semantically relevant then it seems dubious to have a definition for it at S-structure that differs from its LF characterization.

Suppose that derivations start at LF and (3) is valid: move alpha only if necessary. Since a condition like (10) holds for wh-phrases and +WH heads at LF, at this level the wh-phrase and the +WH Inf in (11) both have to be in CP for the required spec- head configuration to be established. No change is forced between LF and S-structure, therefore the only legal S-structure must be the one corresponding to (11) itself.

4 LF

Brody (1987, to appear) suggest that D-structure should be characterized in terms of LF chains as in (13):

(13) (of the set of positions in chains) at D-structure all and only root positions are present

Note that (13) is not meant as an additional stipulation, but rather as an alternative to the characterization of D-structure as "pure representation of thematic structure", ie. the level where all and only theta positions are present. As noted there, it follows from (13) and the projection principle that only root positions of chains can be theta positions, since by the projection principle theta positions have to be present at every level. Since D-structure contains only chain-roots, only these can be thematic. Thus it follows that movement can land

---

Hornstein and Sportiche (1981)).

It is not clear however if the result is sufficiently general. The prohibition against LF movement starting from an A'- position is relevant also to non-wh quantifiers: S-structure scope is fixed when the quantifiers involved are in A'-position, cf. Williams (1988). But the condition in (10) in the text does not seem easily generalizable to non-wh quantifiers since these in general do not require a matching head, in fact non-wh quantifiers usually are not associated with any head at all by virtue of being quantifiers.

[7]Refinements of the principle (14) in the text proposed by Rizzi are irrelevant here.

only in thematic positions. (13) allows chains with multiple thematic roots, as appears necessary in parasitic gap structures.

In Brody (to appear) a definition of the concept of argument in the sense relevant for the theta criterion is proposed that holds at LF. At the same time it is argued there that the theta criterion holds at D-structure and that the definition of argument at LF is relevant for this D-structure theta criterion. These proposals are natural in a framework where LF is the basic level of the derivation and therefore provide further indirect evidence for this position.

References

Aoun, J. N. Hornstein and D. Sportiche (1981) Aspects of Wide Scope Quantification, *Journal of Linguistic Research*, 1.

Brody, M. (1987) On 'Chomsky's Knowledge of Language', *Mind and Language*, 2.

Brody, M. (to appear) Case Theory and Argumenthood ms. UCL.

Chomsky, N. (1985) *Knowledge of Language: Its Nature, Origin, and Use*, New York: Praeger.

Chomsky, N. (1989) Some Notes on the Economy of Derivation and Representation, ms. MIT.

Huang, J. (1982) *Logical relations in Chinese and the Theory of Grammar*, Ph.D. dissertation, MIT.

Pesetsky, D. (1987) Wh-in-situ: Movement and Unselective Binding, in E. Reuland and A. ter Meulen eds. *The Representation of Indefinites*, Cambridge: MIT Press.

Pesetsky, D. (1989) Language-particular Processes and the Earliness Principle, ms. MIT.

Pollock, J-Y. (1989) Verb Movement, UG and the Structure of IP, *Linguistic Inquiry*, 20.3.

Rizzi, L. (1991) Residual Verb Second and the Wh-Criterion, ms. Université de Genève.

Williams, E. (1988) Is LF distinct from S-structure? A Reply to May, *Linguistic Inquiry*, 19.7.