Minimalism, move, and the internal subject hypothesis

José Luis González-Escribano

If CHL is optimal, why should it use costly Move-based feature-checking when a local Merge-based option is available? Why should a 'lazy' CHL bother to activate feature-clusters like Tns, available on verbs, check that FCs and verbs match, and finally remove FCs to satisfy FIP? Why should it furthermore copy-move DPs into the specifiers of FCs, satisfy itself that features match, and then remove all the copies but one?

In MG such analyses result from a view of simplicity that induces biuniqueness between properties and positions, forces categories to be represented by 'chains', and therefore requires a polistratal grammar.

In this primarily conceptual paper, I argue, following CG-HPSG work, that the ISH is conceptually misguided, that FCs are uneconomic computational artifacts, that Move/Attract is as redundant as its triggers, or the deletions it causes, and that dropping them yields a neater theory of CHL.

1. Introduction

Conceptual beauty has always been a priority in Chomskyan linguistics, as in all serious scientific work. Indeed, it is the underlying idea that the world is beautiful in a Platonic sense that has led Chomsky to assume, just as a working hypothesis, but rather explicitly in recent times (see Chomsky,1993:2, 1995a:389-90), that the Computational Component of Human Language (henceforth CHL) is an optimal system.

Certain prominent features of the design of Minimalist Grammar (MG, henceforth), however, are not optimal in any obvious sense, and should in my view be reconsidered. The one I will subject to scrutiny in this paper is its 'multistratal' analysis of clausal processes in terms of (several versions of) the Internal Subject Hypothesis (henceforth, ISH), the Move/Attract metaphor, and functional categories like 'light v', as in Chomsky (1995b), Pr in Bowers (1993), Tr in Collins (1997) and T, AGR, Polarity, Modality in practically everybody's analysis. My starting point is that,

prima facie, a monostratal Merge-based analysis, as assumed in much post-Chomskyan non-transformational grammar (e.g., GPSG, HPSG, Categorial Grammar), is more 'minimalistic' (see Kitahara [1997] on the relative complexity of Merge and Move), and deserves being carefully explored.

The immediate issues in need of consideration, therefore, are a) what Move/Attract does for us, and whether it is a sufficiently well-motivated expedient in MG, and, more generally, as part of the CHL of UG, and b) whether ISH and the various functional categories T, AGR, Tr, v, etc. it rests on yield an optimal analysis in this area of grammar.

As a result of that conceptual scrutiny, I will challenge certain core asumptions of current MG, such as the ISH, the existence of 'functional' categories like T and AGR, or Greed-driven subject and verb raising, and will sketch a more parsimonious alternative in which clauses are just saturated VPs, subjects and predicates select each other in situ and do not move at all, Move is dropped from the grammar, and a generalized version of feature-checking (Generalized Matching Principle) is assumed to monitor Merge.

2. Why move?

Ultimately, Move has always existed in transformational grammar because, assuming the non-obvious hypothesis that grammatical functions are configurationally based and a certain view of simplicity that crucially assumes biuniqueness between grammatical properties and structural positions (see McCloskey,1997), the simplest structures on which such functions can be defined and those that can be assigned to the sequences actually input to the articulatory-perceptual and conceptual-intentional devices happen to be non-isomorphic in marked cases.

Hence, some conciliatory principle is needed, and Move is an intuitively attractive metaphor, although other possibilities exist, e.g. 'Slash' or 'Cancellation', as in the GPSG-HPSG-CG tradition (see Gazdar et al.,1985; Pollard & Sag, 1987, 1994; Steedman,1993, 1996) the 'Visitor' relation of Hudson's WG, as in Hudson (1996), simple unification via functional equations, as in LFG (see Kaplan & Bresnan [1982] or Kaplan & Zaenen,1995[1989]), or the notion of 'chain' in such monostratal models as Koster's (1987) or Brody (1995).

Obviously, different grammatical and discourse functions certain syntactic elements accumulate sometimes depend on incompatible structural configurations. Thus, in English, the optimal structure for the definition of 'object' of a verb, for example, is assumed to be [S [VO]],

which captures the essential subject-object asymmetry as well as the preferred surface order. The function of 'topic', on the other hand, is best defined on the structure [Top [S]] (or variants thereof such as Rizzi's (1997) 'split Comp system' analysis). Hence, if an item is simultaneously object and topic, as is 'cheese' in expressions like 'Cheese, I hate', 'Cheese, she knows I hate', etc. no single monostratal representation can reconcile the two optimal structures required. The same occurs when a wh-item is both 'focus' of a question and, say, object of the verb, like 'what' in 'What did you do?', 'What did you say you did?', etc.

In such a plight, the grammarian is bound to choose between a disjunctive, multiclausal definition of object in a monostratal theory and a monoclausal definition in a polistratal one. The latter strategy is arguably more satisfactory in general, but requires a principle that systematically relates strata, and in the TGG tradition that principle is Move/Attract.

In addition to such relatively uncontroversial cases of 'displaced' constituents, P&P/MG syntacticians have posited other rather less obvious instances of movement. For example, as we said, by assuming ISH and a host of functional categories, they are forced to postulate the ascension of the subject and the verb (or of certain feature-sets within such elements, as in Chomsky's 1995b 'chapter 4') in order to pick up/check off their inflectional features, as in Chomsky & Lasnik (1993) (or rather those of the host FCs in the more recent versions). Even more controversially, perhaps, by assuming the theory-internally optimal, but non-obvious, hypotheses that a) C-I systems also compute P-markers and b) quantifier scope is configurationally based, they are led to postulate even 'invisible' (LF) movement (see Hornstein,1995). Thus, in order to account for the ambiguity of 'Everybody here speaks two languages' in structural terms, they have resorted to the expedient of assigning to that string the two (simplified) LF structures a.[Everybody here [Two languages [x speaks y]]] and b.[Two languages [everybody here] [x speaks y]], both non-isomorphic with the canonical [S [VO]] structure, which in this case coincides with the preferred analysis of that expression at the P interface (assuming the latter to minimally satisfy left-to-right order of articulation).

In sum, P&P Move is conceptually dependent on 1) the methodological ideal of maximum simplicity of statement, 2) the notion that CHL is an optimal computing device, 3) the structure-dependence hypothesis, 4) a certain biuniqueness assumption according to which each separate property must be associated with a distinct structural position (see McCloskey,1997) for a very clear statement in this respect), and 5) highly specific, though certainly defensible, executional preferences concerning the nature of structures themselves.

MG adopts the same overall analytic and descriptive strategy as the P&P version in this respect, except that, contrary to earlier Move Alpha, minimalist Move/Attract rests on feature-checking and is no longer free, but driven by Greed or, in later versions, by the notion of 'enlightened self-interest'. Technically, MG forces movement to exist via the assumption that feature-checking a) is required - minimally at L, but also at Spell Out for certain designated features, b) is structurally based, and above all c) is non-local. Thus, assuming ±WH to be available on C only, the 'strong' +WH feature of interrogative Comps forces interrogative wh-phrases to become specifiers of suitable Comps before Spell Out; assuming nominative Case to be strong and available only on T (or, equivalently, T to have a strong EPP feature that must be checked off), the DP subject must rise overtly and become a specifier of T; similarly, under essentially Larsonian assumptions and Hale & Keyser's (1993) views on rôle assignment, a transitive verb must rise to Tr or v (a variant of feature-checking) in order to assign the Agent role to its external argument; finally, via the auxiliary hypothesis that auxiliaries are invisible at L, as tense and agreement features of verbs must also be checked off, such features on an auxiliary force it to move overtly and adjoin to the Tns-Agr complex, producing the desired surface order of tensed 'have' and 'be' with respect to negation and VP adverbs.

These are the core cases of overt movement in English. In addition, the relevant feature sets of the object and the lexical verb if inflected for tense and agreement must ascend to AGRo, after Spell Out. All these processes follow from Greed/Enlightened Self-Interest and the assumption that inflectional features are computationally relevant and must have been checked off at the L interface.

However, not all observed cases of displacement of constituents fall out as straightforwardly. What forces such overt processes as Topicalization, XP-Preposing, Though-Movement, Extraposition, and Heavy-NP Shift, or covert ones like Operator and Q-Raising, under this inflectionally inspired account is not obvious, since in many languages (e.g., English) no inflectional features seem to be involved in at least some of those cases, so a broader class of computationally relevant triggers, presumably including discourse functional features like +topic, +focus, etc., must eventually be set up if the program is to be carried out in full (see Rizzi [1997] for a plausible extension of the set of triggers along these lines). Let's assume this can be done, though, and proceed to examine certain conceptual issues that arise in the pursuit of the minimalist strategy as designed in such standard works as Chomsky's 'chapter 4' in Chomsky (1995b).

3. CONCEPTUAL TROUBLE

At least in some of the key instances (Case, tense, and agreement features) the non-local feature-checking implied in the above analysis of the clause is a conceptually suspect expedient in several respects.

On the one hand, it calls for the existence of questionable lexical objects like [Tense:+,Past:-], [Person:3,Plural:-], and much more complex grammatical feature-bundles in languages with richer inflection than English. Such feature-bundles are currently assumed to be selected from the lexicon as components of the numerations that start derivations (cf. Chomsky,1995b), although it is by no means clear whether this is a one-step process, as the ideas of 'numeration' and 'reference set' seem to imply, or a 'gradual' multi-step one, as the notion of general availability and uniformity of syntactic operations along the Lexicon->L computation requires (see Brody,1995), Collins,1997:89-93).

Be that as it may, as far as I know, no clear source has ever been offered for such functional items. The fact that they are semantically otiose (in fact they must be eliminated to avoid violations of FIP) and that their composition is context-sensitively determined makes them highly implausible as independent lexical entries, if only on conceptual grounds.

A more natural view is to understand them as, in the most favourable cases, mere 'copies' of feature-complexes that must be generated in verbs and nouns anyway, which is where, following the strict lexicalist assumptions of MG, they belong. The qualification 'in the most favourable cases' is needed insofar as, in principle, it is possible for a random mechanism like CHL to create mismatches by systematically activating the wrong inflectional complexes, causing the derivation to crash. Needless to say, such randomness of lexical choice makes CHL liable to serious psycholinguistic objections that may be neutralized only if an omniscient supervisor is somehow invoked to monitor the lexical selection process. Of course, that may not be the task of CHL itself, but that of some other interacting system (say, pragmatic competence?), but, even in that case, there is no reason to assume that CHL is not inherently constrained in such a way that Select is not random. The problem is a general one, though, and I shall not go into it here. For our present concerns, it suffices to point out that the obvious way to avoid this particular problem in the case at hand is to eliminate the possibility of mismatch by dispensing with artificial inflectional nodes.

Notice that if T and AGR are understood as copies of feature-bundles independently available on verbs and nouns, they are not only questionable lexical items, but imply a copying operation which is not

Move (cf. the 'copy theory of movement') and, above all, turn into plainly redundant computational contrivances.

Indeed, such functional sites pose a problem of redundancy comparable to that of early category-specific PS rules with respect to lexical subcategorization patterns, i.e., their range will necessarily be a mere summary of inflectional variation on verbs and nouns. In this respect, we may invoke a further analogical argument from computational minimality. As Chomsky himself says (1995a:404), structure-preserving substitutions were artificial and redundant, and have been duly disposed of in current MG, because the target-position had to be available before the actual movement took place, an impossible condition to fulfil once D-S is eliminated. If that argument is found compelling, then the same Occamian reasoning should be extended to abstract functional sites like T and AGR. Notice that, to the extent that the exponents of functional categories cease to be lexical items with semantic and phonetic content, they turn into mere computational artifacts (partial projections), and if these have been considered illegitimate in the case of projected X-bar information (see Chomsky,1995a), as a matter of principle they should be dropped in this case, too. After all, parsimony is the key asset of MG.

Finally, from a broader minimalist perspective, that of any reasonable interpretation of economy, such items automatically induce the question of why should an otherwise 'lazy' mechanism proceed to make/insert redundant copies of features immediately available, scatter them about its working area, and then laboriously reassociate them with the originals via computationally costly movement/attraction operations?. The only (metaphorical) answer that occurs to me is that the original 'signal' provided by the verb's morphological features may not be 'robust' enough to guide the computation process in the outer layers of expansion and must therefore be 'echoed' at certain key intervals (i.e., when vital links must be established) until the process is complete, but nobody, as far as I know, has explicitly pursued that line of thought. The standard reasoning on the topic rather invokes the idea that the compact morphological realization of the various features instantiated on the verb hides important hierarchical (e.g., scopal) properties that must be explicit at L, but the only evidence brought to bear in this respect comes from surface word order. Nobody, as far as I know, has invoked deep principles to account for the disposition of AGR, TNS, Pol, Modality, ASP, etc., which, incidentally, is assumed to differ cross-linguistically. Thus, pending further evidence concerning the scopal relations between functional categories at UG, if such evidence can be found, this argument has no force.

In sum, in conceptual terms, and strictly in keeping with the aesthetic principles underlying Chomsky's philosophy of science, it is more natural to drop T and AGR and assume that inflectional features are licensed (along with thematic, selectional, and subcategorization ones, cf. infra) in situ. Needless to say, if this unorthodox assumption is adopted, one of the key sources of theory-internally imposed movement disappears from MG, a fact that should be taken into account on evaluating the issue. In what follows I shall try to show step by step that this alternative design is technically possible, more parsimonious, minimalist in spirit, and just as effective in empirical terms.

4. DROPPING GOVERNMENT-BASED THETA-MARKING AND THE ISH

In the last decade, mostly aesthetic locality considerations related to Theta Theory have made many researchers endorse various versions of ISH (cf. Koopman & Sportiche,1985; Fukui & Speas,1986; Sportiche,1988; González-Escribano,1989; Pollock, 1989; Koopman & Sportiche,1991; Chomsky,1991, 1993, 1995a,b; Chomsky & Lasnik, 1993; McCloskey,1997; etc.). The crucial idea is that thematic processes operate under government, so that arguments of a head are to be realized in positions contained within the head's maximal projection. As there is ample evidence from semantic selection that the subject is an argument of the head X, it follows that it must be internal to the XP.

On the other hand, as claimed in Chomsky (1981), Marantz (1984), and ever since, there is some evidence that the subject-argument of an XP is compositionally theta-marked by the lexical head X and the remaining complements. Thus, under the assumption that predicates and arguments must c-command each other, it follows that the subject must asymmetrically c-command all the other arguments, a hypothesis that receives strong support from the facts of Binding Theory.

Hence, if only specifier and complement are potential A and Theta positions, and all other arguments are realized as subcategorized complements dominated by the lexical projection X', as in the standard version of X-bar theory, the lowest A-position available to the subject is Spec of X, Spec of V (or 'v') in the core case. That has, indeed, become the usual view associated to the now more or less uncontroversial ISH. Larson (1988), however, generates objects in Spec of the lower V and posits an extra VP shell to have a higher Spec slot to lodge the subject in, but his subject still sits in Spec of (the higher) V at initial structure. Koopman & Sportiche (1991), in their turn, generate subjects not in Spec of V, but Chomsky-adjoined to VP, although still under a V projection,

and there are several other variants of the same basic idea, but we may disregard such details here.

It must be pointed out, though, that the conclusion that the subject occupies Spec of V/'v' is somewhat infradetermined in P&P theory, because it rests on the unmotivated idea that only X'-internal slots and Spec of X are A-positions (and possibly theta positions). However, under one of the most influential theories of government (Aoun & Sportiche,1983), any other slot dominated by XP might be equally available for theta-marking, (and in fact must be, if it is assumed that adjuncts are also Theta-marked). Thus, a less stipulative and more encompassing view of A-positions and Theta-marking would immediately make various other slots available, e.g., the subject might be adjoined to various X' nodes, or even hang from XP without occupying Spec (for example, it could be projected in postnuclear position, as argued in González-Escribano,1989).

Be that as it may, it is by now fairly obvious that any attempt to handle all propositional structures (i.e., both full clauses and small ones, cf. Stowell,1981, 1983) in a uniform way along the standard ISH lines leads to inconsistency. The problem is that, as pointed out, among others, by Bowers (1993) or Pollard & Sag (1994), the subject of small clauses cannot possibly occupy the Spec of its predicate, because that position must often be reserved for other specifiers, as in 'I consider John my best friend', where 'my' must occupy the Spec of D, under Abney's (1987) now standard analysis. Yet, granted the assumptions underlying the small clause analysis, it is reasonable to maintain that 'John' is the subject of the predicate DP 'my best friend', so 'John' cannot be in Spec of DP if 'my' is.

Thus, it seems as though the locality requirements for Theta marking may have to be weakened somewhat to include positions outside XP, which implies abandoning government-based views of subject Theta-marking and the strong version of ISH. Indeed, Koopman and Sportiche (1991) do weaken ISH by lodging the subject under an extra projection of V (Vn) outside the X-bar system, and Bowers (1993) definitely returns to a variant of traditional P&P analyses like Williams's (1980, 1994), Stowell's (1981), or Chomsky's (1981), by projecting it as the Specifier of an extra functional category Pr governing (and therefore external to) the XP predicate that Theta-marks it.

So much for the state of the art, aptly summarized in McCloskey (1997). The deep question at stake here, however, is whether the aesthetic motivations underlying ISH are conceptually sound, i.e., after all, why should we expect all arguments of a head to be realized internal to its maximal projection?.

On its aesthetic side, ISH rests on considerations of a) symmetry, i.e., if other arguments are Theta-marked under government and have to be realized inside XP, the subject should be no exception, and b) thematic autonomy of phrases: by requiring that the Projection Principle be satisfied within the boundaries of XPs we make maximal projections self-contained and presumably autonomous in thematic and referential terms.

The symmetry consideration is compelling, in principle, but the thematic and referential autonomy of phrases, although obviously attractive in certain cases (e. g. DPs and CPs often refer to discrete entities and states of affairs, respectively) cannot be generally subscribed, because it ignores an essential aspect of natural language architecture, namely that linguistic expressions must usually be *constructed with* others.

As soon as the constructional aspect of language is taken into account, though, maximality and thematic saturation must go apart, and the intuitive appeal of ISH vanishes, for the constructive relation *par excellence* is, of course, that of a predicate towards its arguments, and, if phrases are identified with fully saturated predicates, it follows that they cannot be further constructed with anything. Syntax, in other words, ultimately depends on the existence of two classes of terms, saturated (arguments) and non-saturated (predicates, CG's 'functor categories').

The idea underlying ISH, in other words, is that maximality implies thematic saturation. That hypothesis works fine for DP and CP arguments but not for AP, PP or CP adjuncts, DP predicates, or PP arguments, for that matter. Notice that adjectival and prepositional adjuncts (of N', say) can only be licensed under the assumption that they are semantically unsaturated one-place predicates taking their respective heads (or, at any rate, some feature or variable in them) as the saturating argument (see Higginbotham,1985). Hence, if, under a government-based theory of Theta-marking, maximality implies full satisfaction of argument structure, they cannot be maximal projections unless they contain perfectly ad hoc PROs coindexed with the head the adjuncts modify, a dubious move, since, a) under Aoun & Sportiche's definition of government, such PROs would be governed and violate the PRO Theorem, and b) certain 'modifiers' of A, P, etc., such as DegP, intensifier 'right', etc., have often been considered as canonical specifiers of A, Adv and P (e.g., by Jackendoff,1977) which would render them incompatible with PRO. If such modifiers are adjuncts, of course, the spec position would still remain available to PRO, but argument a) stands, and, on the other hand, examples like 'I consider John my best friend' cannot be accounted for.

In sum, it is rather plausible that such adjectival, adverbial and prepositional adjuncts are just unsaturated predicates (as analyzed in

Categorial Grammar), i.e., X' categories in X-bar terms. If this reasoning is correct, in other words, Theta-marking based on government and ISH will have to be abandoned: an X does not satisfy all its arguments internal to its XP, and it seems as if we must come back full circle to subjects as 'external' arguments and ultimately to some version of Williams's (1980, 1994) Theory of Predication.

Indeed, in MG, specially if Theta-marking ceases to depend on (lexical) government, ISH is completely unmotivated, and can best be considered as just a relic of earlier theorizing, a ghost of D-structure. Notice that in Chomsky (1995a,b) maximality is dynamically assigned to terms that do not project further, independently of their status with respect to thematic saturation, i.e., a term may be [+Maximal] without being thematically saturated, or equivalently, a head X must be allowed to Theta-mark one (or more, cf. González-Escribano [1997a] on 'Discontinuous APs' like 'a similar car to mine') of its arguments outside X [+Maximal].

It follows that there is no reason why the subject should be expected to occur in Spec V (assuming such a position can be defined), although it is essential to the saturation process (Projection Principle, Theta Criterion, Predication Principle, EPP,...) that a subject should occur at L somewHERE on the P-marker. But, if the preceding reasoning is accepted, subjects do NOT have to move in order to reach a position c-commanding the rest of the predicate and agreeing with the inflected verb. That may as well be their initial (and only) position (see Dowty [1982] or HPSG work like Pollard & Sag,1994).

Notice, in support of that analysis, that the minimalist account does in no way require that the subject be first projected precisely in Spec of V. On the contrary, according to strict minimalist assumptions, if V has no feature-checking potential in situ and the morphological features that need checking must, by stipulation, be checked in Spec of T, or Spec of AGRs, Spec of V should NEVER attract the subject.

In earlier work, Chomsky himself assumes that Spec of T may be left unprojected, since the subject's Case and AGR features are jointly checked off at Spec of AGRs, and, by analogy, we may assume the same to be true of the Spec of untensed auxiliaries when not filled by 'floating quantifiers' (see Sportiche,1988). The present account simply goes a step further: with the disappearance of X-bar theory under Chomsky's (1995a) 'bare phrase structure' hypothesis, the Spec position loses its identity and ceases to exist as such (see also Kayne [1994] on the equivalence between specifiers and adjuncts). Of course, I continue to assume that heads with feature-checking potential 'attract' constituents, and merge with various specifiers (adverbs, 'not', genitives, etc.), but if the feature-

checking potential is absent from the head, as implied for the lexical verb by non-local views on checking, just nothing will be projected in Spec of V.

In sum: under strict minimalist assumptions, feature-checking heads in the I system are limited to T, which can jointly check Case and agreement features. It follows that only T, not V, should 'attract' DPs. Whether such 'attraction' has to be construed in terms of 'Move' (feature-copying) or just in terms of Merge is quite a different issue to which we may now turn.

5. Dropping Movement: Towards a monostratal approach

As pointed out above, movement-based polistratal analyses arise from assumptions concerning simplicity of the theta marking rules and depend on the existence of functional categories, which, as we saw, create substantial conceptual difficulties if understood as lexical items and appear to be equally objectionable as computational artifacts. On the other hand, as we have seen in §4, ISH is empirically dubious, possibly incoherent in conceptual terms, and technically not more, but less, akin to minimalism than its Predication-based alternative.

However, at the same time, as seen from outside MG, non-local, Move-mediated feature-checking is obviously a computational artifact. There is no conceptual reason why a fully inflected verb should be declared unable to establish agreement (or any other interdependence, for that matter) with its subject or object in situ (as required in standard analyses within such varied traditions as Word Grammar, GPSG, LFG, HPSG or CG). On the contrary, computationally, it is more economic (see Kitahara [1997] for arguments), and therefore more sensible in minimalist terms, to let the verb license all its features locally. That strategy, nevertheless, does not necessarily lead to ISH. It may also lead to a version of Predication Theory dispensing with the extra structure built around T, AGR, etc., along with the ad hoc movements it induces (or, equivalently, the chains postulated by Brody,1995). What would seem to stand in the way, of course, is the possible complication of the Theta Marking rules which would be needed.

Technically speaking, however, there is no real problem, as proved by the well-known standard 'gapless' analyses of discontinuity in revised HPSG (see Pollard & Sag,1994: chapter 9), and CG (see Steedman,1993, 1996). We may initially assume, as in CG for the sake of concreteness, that when their specifications 'match' (in a technical sense defined below), unsaturated predicates merge with one complement at a time to form

partially saturated binary-branching predicates, and that 'subjects' finally merge with such projections, agree with them, and thematically saturate their last argument (see Brody,1982). That yields, for example, simple S,[V,O] structures, order parameterized.

However, a partially saturated verbal projection V, with appropriate morphology, may also merge with, say, an aspectual predicate 'be', which projects, and, in Chomsky's terminology, causes the reevaluation of V* as [+Maximal], although still unsaturated. The corresponding CG process could just be ordinary (Forward) Functional Composition (see Wood [1993] or Steedman,1993,1996). The 'be' projection, in turn, (again, assuming adequate morphology) may merge with a second aspectual predicate 'have', or with a modal M, which thereby satisfy their subcategorization features, project 'have' or M phrases, as the case may be, cause the reevaluation of the 'be' projection as [+Maximal] (and unsaturated), and leave the resulting 'have' or M projections still unsaturated through inheritance of an undischarged argument from one of their constituents. Eventually a properly inflected predicate will merge with an agreeing DP and become a thematically saturated expression (possibly a root node) or, alternatively, the saturated predicate will become the argument of (and automatically saturate) a monadic predicate (say a sentence-level adjunct), and so on.

The technical implementation of this intuitive idea is quite straightforward and echoes the Slash convention of GPSG-HPSG or argument cancellation in CG as well as various 'linking theories' within the P&P tradition: it suffices to, say, asterisk argument variables in a head's theta grid only as appropriate XPs 'interpret' them by joining the projection via successive mergers. By way of illustration, in 'John has been doing some work in his studio', 'doing some work' is [+Maximal] (since 'been' projects) and yet unsaturated (i.e., its thematic representation is [do <AG,TH*>]), so is 'doing some work in his studio', which results from the V' projection saturating the external argument of the preposition 'in' (a P', and yet [+Maximal] in Chomsky's terms), and so are 'been doing some work' and 'has been doing some work', since the auxiliaries 'be', and 'have' satisfy their own subcategorization requirements (i.e., [__V+ing] and [__V+en], respectively) via the corresponding mergers, but cannot discharge the AG role of the full verb 'do'. Only when 'John' merges with the higher predicate does the latter become saturated (i.e., thematically 'do <AG*,TH*>') and therefore, by definition, the resulting phrase ceases to be a predicate.

More perspicuously perhaps, we may construe a head's feature specification, and particularly its Theta Grid, as a set of functional [R(e):v]

pairs, where R is a theta role, 'e' is an event/entity Davidsonian variable, and 'v' an index (e.g., [AG (e):i]). We may further assume that part of the feature specification of argument XPs is a [Ref(erence)(xp): i] attribute, and let the value of [Ref(xp): i] unify with, and bind, the variable 'i'. Each successful merger binds one of such variables on the head's Theta Grid until no variables remain uninterpreted (Full Saturation, Projection Principle). The Theta Criterion, on the other hand, automatically follows from the, in LFG terms, Functional Uniqueness requirement such [R(e):v] pairs imply, by definition.

Adjuncts, being monadic predicates (see Higginbotham (1985) and all the above-mentioned CG tradition), may be handled in essentially the same way. We may assume their thematic specification to contain functional pairs such as [Time(e): t], [Place (e): p], [Manner (e): m], [Frequency(e): f], etc., and let the 'e' variable correspond to the adjunct's external argument (i.e., the syntactic head N, V, etc.) and the values 't', 'p', 'm', etc. to the X' adjunct proper (or its head's internal argument, as in the case of prepositional adjuncts or A's containing complements).

Notice that, under this strictly relational approach to thematic notions, the Davidsonian event argument coincides with the whole functional structure 'e', i.e., there is no 'extra' event argument. Reference to a specific real world event is achieved via the 'e' variable that figures as the internal argument of R(e):i equations.

As to the technicalities of the merging process, on both its 'horizontal' (head-complement/subject) and its 'vertical' (head-mother) side, it is not crucial to settle such details here, but there are several possibilities which merit consideration. One, of course, is unification, as in Kaplan & Bresnan's LFG, Pollard & Sag's revised HPSG, Kay's (1985) Functional Unification Grammar, or Karttunen's (1989) Categorial Unification Grammar. Unification seems appropriate to handle 'vertical' feature-passing, as a replacement for such principles as Percolation, or the Head Feature Convention of GPSG-HPSG, but not quite so adequate to handle the matching implied by Merge at the horizontal level, which in CG is based on 'multiplication' (set intersection).

The feature-checking approach usual in standard MG (e.g., Chomsky,1995b) is feature-correspondence based on the actual *duplication* of features on heads and nonheads, and, to that extent, it is redundant and fails to capture the 'complementarity' between predicates and their arguments.

The approach I find conceptually more satisfying (see González-Escribano,1997b), much in the spirit of categorial and some dependency grammars, is based on the idea that the items that merge are not (partly)

identical; they are 'different', but strictly *complementary*, i.e., a transitive verb 'selects' an accusative object and a tensed verb 'selects' a nominative subject, for example, but we do not say that the transitive verb 'carries' an accusative case feature that it must assign to an appropriate DP, nor that a tensed verb 'carries' a nominative case feature that must be checked off against an appropriate nominative DP subject, etc. That would imply duplication of the same feature-sets and ultimately deletion if FIP is to be satisfied. On the contrary, the alternative defended here implies neither duplication nor deletion. Only one token of each feature will enter the derivation (see González-Escribano,1997b) for a systematic exploration of this approach).

Implementational matters aside, the theoretical picture that emerges, in other words, is as follows: clauses are just saturated XPs (VPs in the core case); T, AGR, Polarity, Modality, etc. do not exist as independent terms, but only as grammatical feature complexes of real lexical items (full verbs or auxiliaries, in the case of English); feature-checking is strictly local and operative when terms with suitable feature complexes merge (i.e., under strict mutual c-command); subjects, in particular, merge successfully with the predicates that select them for Case, person, and number, and cause crashes otherwise; subject movement is not a grammatical process, since the subject does not pre-exist in any lower position, but is directly attracted by, and merged with, the highest verb, be it a full verb or an auxiliary; finally, there being no independent T node, verb raising to T does not exist either, and therefore a nice result follows: there is no need to relax the general procedure that reevaluates the PS status of nodes as [±Maximal] to avoid licensing an heterogeneous chain (see Chomsky, 1995a:409).

In order to account for the subjects of small clauses and non-finite clauses uniformly under this approach, we may assume that the higher verb (or the complementizers 'for' or 'with') select and mark the SC, infinitive or gerund clause, so that its subject in fact agrees with its predicate. It follows that a lexical subject becomes illegitimate in subject position if the clause is not selected by a higher verb or Comp, or if a non-marking Comp (i.e., empty C) intervenes. In such circumstances only a controlled PRO may appear as subject, as desired.

In sum: granted some feature-passing convention, subjects can be monostratally generated by an appropriately constrained version of Merge. The reasoning required, in general, is a close translation into our terminology above of that used to motivate Slash/Cancellation in GPSG, HPSG, or CG, so we shall not dwell upon it here any further. The point is just that it is possible to replace Move with Merge at no empirical cost,

and at some conceptual advantage, and, indeed, not only in the special case of purported subject and verb movement we have discussed here, but in general, since the same approach can be straightforwardly extended to WH-phrases and other displaced constituents.

6. ACCOUNTING FOR SURFACE WORD ORDER

The spectacular proliferation of functional categories in recent years was largely triggered by two factors, one theory-internal and one empirical. The former derives from the structure-dependence hypothesis, as combined with a certain notion of simplicity, the idea - also underlying ISH, as we saw above - that properties and structural positions must be biuniquely related. Thus, the compact morphological realization of the various features usually instantiated on verbs is taken to hide important hierarchical (scopal) properties that must be represented in structural terms at some level, which in the current minimalist model can only be L. The prima facie evidence for that approach differs depending on the languages investigated, being more substantial in richly inflected ones, where the various affixes are often visibly ordered in ways that suggest what their scopal properties must be. In Modern English, on the contrary, direct morphological evidence for the structural approach is weak.

The empirical factor, on the other hand, has been the need to account for surface order. What inflectional nodes like T, Modality, Polarity, and AGR do, in essence, is allow grammarians extra hierarchically ordered positions to accommodate auxiliaries, negation, and various adverbs with respect to each other. In fact, the distribution of adverbs, along with the biuniqueness assumption, has, in part, made up for the absence of direct morphological evidence and allowed for the extrapolation of hypotheses concerning specific functional categories from particular languages to Universal Grammar. Such universal FCs are subject to parametric variation only in such respects as the inventory actually attested in each language and their hierarchical disposition in clause structure.

However, nobody, to my knowledge, has invoked deep principles to account for the relative disposition of Agr, Tns, Modality, and ASP either at UG or in any particular language, and no attempt has been made to make the presumed parameters precise. As such 'attractors' in turn determine the visible position of DPs and verbs, the result is that clause structure and word order have been stipulateD on the basis of surface evidence (as mediated by the biuniqueness hypothesis and the mechanics of verb raising), but never properly EXPLAINED. To that extent, I believe it is fair to say that FCs are being used as mere computational artifacts. In

the absence of deep principles explaining the scopal relations among such categories at UG, the FC approach may be descriptively successful, but is stipulative and has no real explanatory content.

Unfortunately, as regards such deep principles, I have nothing substantial to offer. Apart from conceivable language-particular lexical stipulations (e.g., declaring certain verbs 'defective' on the basis of their combinatory behaviour), nothing in the alternative theory sketched here forces, say, English aspectual 'have' to merge with an aspectual 'be' projection, 'have' projecting, instead of viceversa, or negation to merge with the higher ASPP and not with the lower VP or with ModP, etc. In other words, the standard MG and the alternative sketched here are in the same boat in this respect.

The problem is a general one, as we anticipated above, namely that standard Merge does not impose any conditions on the participant terms. Of course, inappropriate mergers are assumed to eventually cause crashes at L, but no details have been offered concerning what makes mergers unsuccessful, and no device has so far been provided to abort them before they reach the interface. It is clear, though, that current views limiting computational relevance to morphological features are much too restrictive (in fact unable to account for all the presumed cases of movement, as we saw), and that, irrespective of whether Move is dispensed with, as advocated in this paper, or not, a generalized Merge-oriented version of feature-checking extending to categorial, thematic and inflectional information is likely to be needed to constrain CHL and monitor its computations process at all stages, beginning by lexical selection. That seems to be the approach taken in Fujita (1996), as quoted by Haegeman (1997), although I have not been able to see that work. In earlier work of mine (González-Escribano,1991,1993), I referred to such a device by the name of 'Generalized Matching Principle' (GMP). A dependency-based version of the same basic idea is explored in some technical detail in González-Escribano (1997b).

In sum, neither the standard minimalist theory nor the alternative designed here reaches explanatory adequacy concerning the deep problem of what guides Merge, although the GMP-based version defended here is more constrained and may avoid the overgeneration problem and the psycholinguistic objections it motivates. However, as regards descriptive adequacy, once the tensed verb, modality, negation, aspectuals and the lexical verb are hierarchically ordered on the basis of surface evidence, the distribution of the subject and the various VP, aspectual, modality, and 'sentential' adverbials may be captured in this approach by imposing appropriate selection conditions on various predicates or

features thereof as efficiently as in a theory resting on functional categories. Finally, as regards conceptual neatness and economy, the present approach can be claimed to be neatly superior to standard MG insofar as the same results are achieved without artificial lexical items, conceptually questionable assumptions like ISH, Move, initial feature-copying, and the subsequent deletions involved if FIP violations are to be avoided.

José Luis González-Escribano
Universidad de Oviedo
Campus de Humanidades, Room 2310
C./ Teniente Alfonso Martínez, s/n,
33011 Oviedo, España
Tel. (34)(8) 5104587/4532
Fax. (34)(8) 5104555
E-mail: escri@airastur.es
escri@pinon.ccu.uniovi.es

REFERENCES

Abney, S. (1987) *The English DP in its Sentential Aspects*. Ph.D. Diss. Cambridge, Mass.: MIT Press.

Aoun, J. and D. Sportiche (1983) "On the formal theory of government", *The Linguistic Review* 2, 211-236.

Bowers, J. (1993) "The syntax of predication". *Linguistic Inquiry* 24, 591-656.

Brody, M. (1995) *Lexico-Logical Form. A Radically Minimalist Theory*. Cambridge, Mass.: MIT Press.

Chomsky, N. (1981) *Lectures on Government and Binding*. Dordrecht: Foris.

Chomsky, N. (1991) "Some notes on economy of derivation and representation. In R. Freidin (ed.) *Principles and Parameters in Comparative Grammar*. Cambridge, Mass.: MIT Press. 417-454.

Chomsky, N. (1993) "A minimalist program for linguistic theory". In K. Hale and S. J. Keyser (eds.) *The View from Building 20*. Cambridge, Mass.: MIT Press. 1-52.

Chomsky, N. (1995a) "Bare phrase structure". In G. Webelhuth (ed.) *Government and Binding Theory and the Minimalist Program*. Oxford: Blackwell, 385-439.

Chomsky, N. (1995b) *The Minimalist Program*. Cambridge, Mass.: MIT Press.

Chomsky, N. and H. Lasnik (1993) "The theory of principles and parameters". In J. Jacobs et al. (eds.) *Syntax. An international handbook of contemporary research*. Berlin: Walter de Gruyter, 506-569.

Collins, C. (1997) *Local Economy*. Cambridge, Mass.: MIT Press.

Dowty, D. (1982) "Grammatical Relations and Montague Grammar". In P. Jakobson and G. Pullum (eds.) *The Nature of Syntactic Representation*. Dordrecht: Reidel. 79-130.

Fujita K. (1996) "Generalized Attract and Economy of Derivation". Ms.

Fukui, N. and M. Speas (1986) "Specifiers and projections", *MIT Working Papers in Linguistics* 8. 128-172.

Gazdar, G. et al.(1985) *Generalized Phrase Structure Grammar*. Oxford: Blackwell.

González-Escribano, J. L. (1989) "Los efectos de inversión y la estructura subyacente de la oración en inglés", *Revista Española de Lingüística* 20/2 (1990, 329-401.

González-Escribano, J. L. (1991) *Una Teoría de la Oración*. Oviedo: Universidad de Oviedo, Servicio de Publicaciones.

González-Escribano, J. L. (1993) "Natural language syntax: transparency and binary branching". In F. J. Ruiz de Mendoza and C. Cunchillos, eds., *Actas del XV Congreso de AEDEAN*. Logroño: Colegio Universitario de La Rioja. 325-344.

González-Escribano, J. L. (1997a) "Discontinuous APs". Submitted to *Journal of Linguistics*.

González-Escribano, J. L. (1997b) "Conceptocentric Dependency-Based Minimalist Grammar: A Realistic Model of CHL". Unpublished Ms.

Haegeman, L. (1997) (ed.) *Elements of Grammar*. Dordrecht: Kluwer.

Hale, K. and S. J. Keyser (1993) "On argument structure and the lexical expression of syntactic relations". In K. Hale & S. J. Keyser (eds.) *The View from Building 20*. Cambridge, Mass. MIT Press.

Higginbotham, J. (1985) "On Semantics", *Linguistic Inquiry* 16, 547-594.

Hornstein, N. (1995) *Logical Form*. Oxford: Blackwell.

Hudson, R. A. (1996) "Discontinuity". Unpublished paper. www.phon.ucl.ac.uk/home /dick/papers.htm

Jackendoff, R. (1977) *X-Bar Syntax. A Theory of Phrase Structure*. Cambridge, Mass.: MIT Press.

Kaplan, R. and J. Bresnan (1982) "Lexical Functional Grammar: A Formal System for Grammatical Representation". In J. Bresnan (ed.) *The Mental*

Representation of Grammatical Relations. Cambridge, Mass.: MIT Press. 173-281.

Kaplan, R. and A. Zaenen (1989) "Long-distance dependencies, constituent structure and functional uncertainty". In M. R. Baltin and A. S. Kroch (eds.) *Alternative Conceptions of Phrase Structure.* Chicago: The University of Chicago Press.17-42.

Karttunen, L. (1989) "Radical Lexicalism". In M. R. Baltin and A. S. Kroch (eds.) *Alternative Conceptions of Phrase Structure.* Chicago: The University of Chicago Press.43-65.

Kay, M. (1985) "Parsing in functional unification grammar". In D. Dowty, L. Karttunen & A. Zwicky (eds.) *Natural language parsing.* Cambridge: Cambridge University Press. 251-278.

Kayne, R. (1994) *The Antisymmetry of Syntax.* Cambridge, Mass.: MIT Press.

Kitahara, H. (1997). *Elementary Operations and Optimal Derivations.* Cambridge, Mass.: MIT Press.

Koopman, H. and Sportiche, D.(1985) "Theta theory and extraction", *GLOW Newsletter* 14, 57-58.

Koopman, H. and Sportiche, D. (1991). "The position of subjects". *Lingua* 85, 211-258.

Koster, J. (1987) *Domains and Dynasties.* Dordrecht: Foris.

Larson, R. K. (1988) "On the double object construction". *Linguistic Inquiry* 19, 335-391.

Marantz, A. (1984) *On the Nature of Grammatical Relations.* Cambridge, Mass.: MIT Press.

McCloskey, J. (1997) "Subjecthood and Subject Positions". In Haegeman, (ed.), 197-235.

Pollard, C. and Sag, I. A. (1987) *Information-Based Syntax and Semantics* I. Stanford: CSLI.

Pollard, C. and Sag, I. A. (1994) *Head-Driven Phrase Structure Grammar.* Stanford: CSLI and Chicago: The University of Chicago Press.

Pollock, J.-Y. (1989) "Verb movement, universal grammar and the structure of IP", *Linguistic Inquiry* 20, 365-424.

Rizzi, L. (1997) "The Fine Structure of the Left Periphery". In Haegeman (ed.) 281-337.

Sportiche, D. (1988) "A theory of floating quantifiers and its corollaries for constituent structure", *Linguistic Inquiry* 19, 425-449.

Steedman, M. (1993) "Categorial Grammar", *Lingua* 90/3, 221-258.

Steedman, M. (1996) *Surface Structure and Interpretation.* Cambridge, Mass.: MIT Press.

Stowell, T. (1981) *Origins of phrase structure*. Ph. D. Diss. Cambridge, Mass: MIT.

Stowell, T. (1983) "Subjects across categories", *The Linguistic Review* 2, 285-312.

Williams, E. (1980) "Predication", *Linguistic Inquiry* 11, 203-238.

Williams, E. (1994) *Thematic Structure in Syntax*. Cambridge, Mass.: MIT Press.

Wood, M. M. (1993) *Categorial Grammars*. London: Routledge.