

# Left-Dislocation *versus* Topicalisation, a semantic characterisation of the former<sup>1</sup>

JAVIER PÉREZ GUERRA

Universidad de Vigo

## 1. Introduction

The title of this article refers to two grammatical devices which have been discussed in almost every syntactic and/or pragmatic approach to the English language. Such mechanisms have often been included under the general label 'fronting,' by means of which a given speaker can 'topicalise' a specific part of speech by locating it in sentence-initial position. Among the different procedures referred to as fronting, grammarians usually distinguish between 'topicalisation'<sup>2</sup> (abbreviated as TOP henceforth) and 'left-dislocation'<sup>3</sup> (LD), illustrated in (1) and (2), respectively:

- (1) LD segments, I shall investigate.
- (2) LD segments, I shall investigate them.

My aim in this paper is to adduce evidence in favour of: (i) an interpretation of both TOP and LD as central syntactic phenomena to be incorporated in the theory of grammar; (ii) a distinction between TOP and LD, the former being explained in a transformational way, and the latter generated in the base; and (iii) an interpretation of LD as a quasi-predicative relation holding between the constituent generated in the base in initial position and its related counterpart in the sentence. Such a claim is in keeping with the pragmatic consequences of LD.

<sup>1</sup> This research was funded by the Spanish Ministry of Education and Science (DGICYR grant no. PB94-0619); this grant is hereby gratefully acknowledged.

<sup>2</sup> Some scholars make a three-fold distinction of preposing systems in English. For instance, Declerck (1988), following Prince (1981), identifies three different operations: 'focus-movement' (our LD) in [*They bought a dog.*] *Fido they named it*; 'topicalization' in [*Mary John saw yesterday*]; and 'Yiddish-movement' in [*They bought a dog.*] *A goat they should buy [ , their house is so dirty]*. Our topicalisation corresponds to both his 'topicalization' and 'Yiddish-movement,' the former being informatively neutral, and the latter implying contrast. Since my study is mainly concerned with the syntactic behaviour and explanation of sentence-initial constituents, there is no need for a classification based only on semantic-pragmatic factors. Syntactically, both Declerck's 'topicalization' and 'Yiddish-movement' prepose constituents and leave gaps in the sentences, independently of the pragmatic value of such actions.

<sup>3</sup> This concept of LD is equivalent to Halliday's (1967) 'reference' or Givón's (1976) 'topic-shift.'

## 2. *Top and LD as syntactic phenomena*

One of the immediate problems with LD or TOP is whether these linguistic tools are related more to discourse than to syntax. Ochs and Schieffelin (1976: 241), for example, claim that "many of the[se] constructions (...) look more like discourses than sentences," and Hankamer (1974) characterises LD and TOP as discourse-conditioned rules on the basis that neither of them can apply in a discourse-initial sentence in the complete absence of relevant extralinguistic material. Even though I agree with the importance of TOP and LD in discourse analysis in its entirety, their syntactic consequences are too strong to be left unexplained, or, in other words, to be included among those linguistic facts that are relegated to solely the fields of discourse or style. Simply because of the repercussions that LD and TOP bring about as far as binding, insertion of interjections, or backward pronominalisation are concerned (see §3), these mechanisms deserve to be taken into account in every analysis of sentence structure. What is more, both historical investigation and studies on language acquisition or applied linguistics reveal that the notion 'non-subject sentence-starter' is probably more central to language than mere statistical accounts show. Consider in this respect some of the evidence provided by analyses within these two fields:

### (a) *Language acquisition*

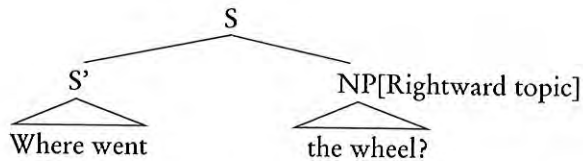
As far as language acquisition studies are concerned, Gruber (1969), for instance, claims that, from child-language evidence, an extra-clausal first-position<sup>4</sup> is absolutely essential in order to explain the sort of questions uttered by children (cf. example 3), declarative sentences of the kind of (4), or the alternation of Case in sentence-initial pronouns, as in (5):

- (3) Where went the wheel?
- (4) Car, it broken.
- (5) {He/ Him} ('s) (a) dog.

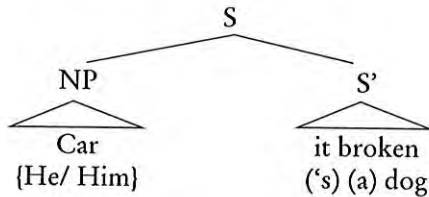
On the one hand, *car* in (4) is not the subject of *broken*, and *him* in (5), even if we wish to analyse it as the subject, is not in the nominative case, as expected in prototypical subjects. On the other hand, since patterns like (3) belong to child-language, together with *Where did the wheel {go/ went}?* or even *Where the wheel went?*, it seems to be the case that the child does not regard *the wheel* as a clear subject when (s)he utters (3) above. Considerations like these led Gruber to analyse the aforementioned examples as in (3') and (4-5')

<sup>4</sup> In Givón's (1983) opinion, the phenomenon of having a constituent in pre-subject position is a reminiscence of a topic-comment word-order. In fact, so-called topic-comment languages —English is not claimed to be a topic-prominent language, but a subject-prominent one (cf. Chafe 1976)— make extensive use of LD and TOP devices.

(3')

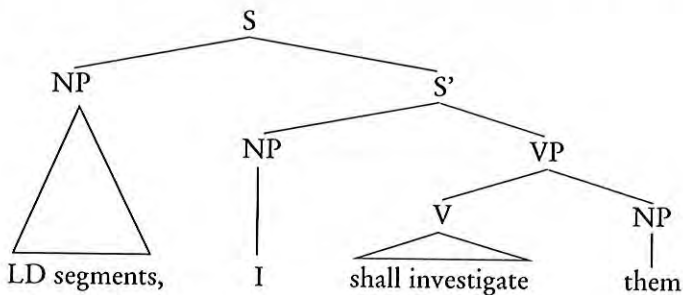


(4-5')



In adult language, this first stage NP [<sub>S</sub> NP VP] is simplified to an [<sub>S</sub> NP VP] pattern, except for TOP, LD and other frontings. Consequently, the phrase-marker in (4-5') is valid for (2) above, as in (6):

(6)



As Bates (1976: 162-63) points out, TOP/LD is, in child language, a passive process, i.e. "by default, as the by-product of focusing and commenting on new and/or important information (...). In adult speech, the T[opic]-C[omment] system becomes far more elaborate, and topicalization and focusing [our LD belongs to the general label of 'focusing'] become active processes."

Givón (1976) advocates a 'de-marking' operation, by means of which, in our (subject-prominent) languages, the adjoined-to-S' position and the NP in S', that is, the subject, are conflated. As a consequence of that process, the 'successful' NP is functionally both topic and subject of the sentence. Givón's suggestion is exemplified in the pair of examples under (7):

(7a) LD: *The man, he<sub>SUBJ</sub> came.*

De-marking: *The man<sub>SUBJ</sub> he-came* [*he-came* coincides with Gruber's Pro-V']  
→ *The man came.*

(7b) LD: *The man, I saw him.*

Rightward movement of the LD constituent: *I saw him, the man.*  
e-marking: *I saw-him the man* → *I saw the man.*

In Geluykens' (1992: 150) words, as far as these studies are concerned,

[t]he interesting consequence is that, from this point of view, LD would be an ontogenetically prior alternative to 'syntactic' subject-predicate constructions. Since child discourse relies much more on the 'pragmatic mode' than on the 'syntactic' mode (...), this shows LD to be a direct reflection of communicative needs. The same can then be said about LD in (adult) unplanned, informal discourse.

(b) *Historical evidence*

With regard to historical research, in former stages of the English language it was a wide-spread practise to repeat a constituent, either completely, or partially, in the shape of a pronoun. Examples like (8), taken from Barber (1976), are quite common, even in Early Modern English:

(8) The skipping King hee ambled up and downe.

In fact, Quirk et al. (1985) refer to LD, in particular, as 'anticipated identification,' the syntactic relationship which holds between, for instance, *the skipping King* and *hee* in (8) being apposition.

### 3. *Differences and similarities between LD and TOP*

At first sight, LD and TOP are similar linguistic phenomena, since in both cases:

(i) There is a prosodic detachment between the fronted constituent and the clause.

(ii) A segment different from the subject appears in sentence-initial position.

(iii) Both LD and TOP constituents are either 'insertable' in subject position or in the subcategorisation frame of the predicate within the sentence following the pause, or, alternatively, they are semantically connected to some constituent in the sentence. In other words, the fronted segments are either syntactically or semantically compatible with the sentence, the former alternative exemplified in (2) above, the latter in (9):

(9) (As for) fronting devices, I have always liked topicalisation and left-dislocation.

In (9), semantically speaking, *fronting devices* makes reference to *topicalisation and left-dislocation*, even though the fronted phrase cannot be inserted in any slot reserved for the primitive syntactic functions required by the predicate *like*, whether subject or object. On the contrary, in (2), *LD segments* can replace *them*, which is a pronominal copy functioning as object.

As a consequence of the obvious relationship between both devices, some scholars, such as Barcelona (1988), do not make an explicit distinction between LD and TOP. Also, from a typological point of view, Foley and van Valin (1985) assert

that, in certain languages, for instance, Bantu, both LD and TOP are coincident or very close linguistic phenomena.

As far as the differences between LD and TOP are concerned, it is obvious that in constructions showing TOP, a constituent 'moves' to sentence-initial position, leaving a gap in the clause, whereas in LD sentences, there is no gap in the sentence, since the subcategorisation requirements are fulfilled either by 'major' lexical constituents, or by pronominal copies. To give an example, in (1) above, the syntactic requirements of the clause after the comma are not fulfilled, since the predicate *investigate* needs a constituent functioning as its object. Such a constituent is *LD segments*, which, instead of appearing in a post-verbal position, as is the case of unmarked declarative sentences in English, occupies the initial position of the utterance. In the post-comma segment, i.e. *I shall investigate*, a gap is perceived, which will become nonsensical unless the speaker/hearer takes into account the topicalised phrase *LD segments*. In the LD version, namely (2), as already pointed out, no gap is observed after the pause, since the object slot is lexically filled by the pronoun *them*. Consequently, *LD segments* is syntactically redundant and unnecessary. It goes without saying that the general concept of LD supported here combines pure LD and clitic LD since it is based on the existence of a segment related to the one in initial position — see Cinque (1990) for the concept of 'clitic LD.'

The identification of TOP poses no problems, since the constituent in first position has to coincide with the gapped position, which is always internal. In other words, by substituting the fronted segment for the gap, the sentence is complete, no pronominal copies or dummies being superfluous. On the contrary, the analysis of LD is, in principle, more complex, as it covers the two different phenomena outlined in previous paragraphs:

(i) The constituent in sentence-initial position is capable of functioning as an argument within the sentence, as in TOP. However, the clausal position of such an argument is already filled by lexical material, to wit, a pronominal copy of the fronted constituent. This type of LD shows what I call 'syntactic compatibility.' As a consequence of the lexical filling of every argumental position, subjects can undergo LD, as witnessed by (10):

- (10) LD constructions showing syntactic compatibility, they trigger interesting consequences.

The TOP version of (10), in (11), is ungrammatical, since English is a non-*pro*-drop language. In other words, (finite) verbs in English must have subject arguments, which cannot be lexically empty — the 'theta criterion' requires that the subject argument must be assigned to some NP.

- (11) \*LD constructions showing syntactic compatibility, *e* trigger interesting consequences.

In (10), on the contrary, the pronominal copy of *LD constructions showing syntactic compatibility*, i.e. *they*, successfully occupies subject-position, thus corroborating the general non-*pro*-drop principle just mentioned.

(ii) The constituent in sentence-initial position makes reference to some argument within the sentence, although it cannot function as such in its place. The word "reference", then, stands for semantic or pragmatic connection, that is, 'semantic/pragmatic compatibility.' Let us check such remarks in view of examples (12) to (15):

- (12) As for London, Trafalgar Square is nice.
- (13) My work, I'm going crazy.
- (14) My work, I'm crazy about.
- (14') My work, I'm crazy about it.
- (15) As far as I'm concerned, London is nice.

In (12), every argumental position in the clause *Trafalgar Square is nice* is filled, there not being a gap ready to accommodate the fronted constituent (*as for*) *London*. Nevertheless, the sentence is grammatical since *London* and *Trafalgar Square* are obviously compatible. Thus, (12) illustrates the subtype of LD showing semantic/pragmatic compatibility. On the contrary, in (13), the only semantic connection holds between *my* and *I*, the former being a non-head part of a constituent, and thus syntactically peripheral within the phrase *my work*. As a consequence of that, (13) will not be regarded as an LD construction but as a sentence modified by the adjunct or disjunct *my work*. A more 'orthodox' version of (13), namely (14), would, doubtlessly, be regarded as a TOP-structure, *my work* being perfectly insertable in the position of the empty complement of the preposition *about*, as reflected in (16):

- (16) I'm going crazy about *my work*.

Example (14') is completely acceptable, and illustrative of the subtype of LD-structures showing syntactic compatibility. As for (15), the segment *as far as I'm concerned* has nothing to do with any constituent in *London is nice*, and thus the whole construction is rejected as an example of LD.<sup>5</sup> I suggest that the fronted segment should be analysed as an aspectual disjunct modifying the clause.

To sum up, TOP and the syntactic type of LD are clear-cut grammatical concepts, since their identification has been strictly based on the syntactic configuration of the sentence. The semantic type of LD, based on the semantic/pragmatic 'affinity' between the preposed segments and some clausal constituent requires a more strict reformulation. A first approximation to the concept LD could be as follows:

- (17) LD requirements:

A sentence is said to contain a left-dislocated segment if it fulfils (a) and either (b) or (c):

<sup>5</sup> Gelyukens (1992), who is mainly concerned with the interactional process of LD and the consequences it brings about as far as communication is concerned, includes (15) among the examples illustrating quasi-LD.

- (a) A constituent other than the unmarked theme (in the sense of Halliday's 'theme') is in sentence-initial position, and a pause (or comma, in writing) is 'felt' to occur between that segment and the rest of the clause.
- (b) A syntactic relation holds between the preposed segment and another element in the sentence. In other words, a copy (or a referent, in those cases showing backward pronominalisation) of the preposed segment occurs in the sentence.
- (c) A semantic connection holds between the preposed segment and another element in the sentence, in such a way that the preposed and the related element in the sentence share the core semantic features. Alternatively, the element in the sentence is 'part' of the preposed one, which thus functions, informatively speaking, as a proper setting for the former.

The definition in (17) identifies what has traditionally been called LD in the literature. However, since the determination of the semantic connection in (c) is a question of prototypicality and not a +/- distinction, I shall momentarily stick only to conditions (a) and (b). Consequently, (14') will be claimed to illustrate LD in English, whilst the fronted segments in (12), (13) and (15) will be treated, in principle, either as sentence modifiers or as discourse facts with no further syntactic consequences.

Once the semantic/pragmatic subtype has been temporarily excluded, the only difference between LD and TOP is reflected by the opposition [ $\pm$  pronominal copy]. That granted, every segment which undergoes LD is also eligible for TOP, with the exception of subjects, as justified some paragraphs above, and some verbal constituents, for which I will account below. Examples (18) to (31) corroborate such a claim, the (a)-series corresponding to LD, the (b)-ones to TOP, and the information in brackets indicating the function of the fronted constituent:

- (18) a. My girlfriend's brother, he likes beans. [subject]  
b. \*My girlfriend's brother, *e* likes beans.
- (19) a. Beans, Steve likes them. [(direct) object]  
b. Beans, Steve likes.
- (20) a. To Peter, I gave the ball to him. [indirect object]  
b. To Peter, I gave the ball.
- (21) Sure about his future, Dudley is not. [subject complement]
- (22) a. President, they appointed her so. [object complement]  
b. President, they appointed her.
- (23) a. At that building, she looked {at it/ there<sup>6</sup>}. [prepositional complement]  
b. At that building, she looked.
- (24) a. In this cupboard, Steve put the beans {in it/ there}.  
[other VP-complements]  
b. In this cupboard, Steve put the beans.

<sup>6</sup> On the basis that *there*, not a proper pronominal copy, perfectly replaces a place-topic, the parameter [ $\pm$ pronominal copy] could be changed to [ $\pm$ pronominal correlative].

- (25) a. \*Did, I {did buy/ bought} a motorcycle. [INFL]  
 b. \*Did, I buy a motorcycle.
- (26) a. \*Buy, he did a motorcycle. [V]  
 b. Buy, he (did) a motorcycle.
- (27) a. \*Bought a motorcycle, he did (so). [Verb group]  
 b. \*Bought a motorcycle, he.
- (28) a. Buy a motorcycle, he did (so). [VP]  
 b. Buy a motorcycle, he did.
- (29) a. The others, I'm still looking into them. [complement of a preposition]  
 b. The others, I'm still looking into.
- (30) That you wanted it, Tom is sure. [complement of an adjective]
- (31) Whether you are there or not, Tom will not mind the question.  
 [complement of a noun]

On the one hand, from (18) to (24), it follows that most major functions are subject to both LD and TOP, if there is not a general blocking principle, as is the case of topicalised subjects. For analogical reasons, I assume that if APs and complement clauses had an available pronominal copy, they would undergo LD.<sup>7</sup> Since there is no grammatical reduced counterpart of these categories, in (21), (30) and (31) I have simply included the TOP version. On the other hand, (25) to (28) show that in I', only VP (i.e. the lexical verb —minus INFL— and its complements) can undergo LD and TOP. Examples (30) and (31) reveal that complements can be extracted from an NP or an AP. What I have no explanation for is the acceptability of (32) with respect to (33). In (32), the TOP of the adjective *sure* governing the *that*-clause is permitted, whereas, in (33), the extraction of *the question* turns into ungrammaticality:

- (32) Sure Tom is that you wanted it.  
 (33) \*The question, Tom will not mind {it<sub>LD</sub> / Ø<sub>TOP</sub>} whether you are there or not.

Examples (25) to (28) show that, as far as the fronting of verbal constituents is concerned, TOP and LD behave in the same way: INFL, V and verb-groups are not eligible for fronting purposes, whereas the whole VP is. The similarities evidenced by the almost utter coincidence of 'dislocatable' and 'topicalisable' constituents are corroborated in (28), a first scrutiny of which reveals that LD and TOP are extremely close in the case of VP fronting. In other words, *did* is understood as the copy of VP in (28a) and as the lexicalisation of INFL in (28b), no difference meddling in the final outputs.

<sup>7</sup> In fact, as pointed out by Cinque (1990), in those languages in which 'clitic left dislocation' is operative (Italian, and, partially, Catalan, for instance), any maximal phrase can be clitic left-dislocated. In others, as determined by Hernanz and Brucart (1987) or Ojea (1991) in the case of Spanish, the clitic-criterion is not always pertinent. However, instead of referring to other mechanisms like the fronted segment's ability to trigger inversion (Ojea 1991) which, in my opinion, may be too language-particular, I will assume that, unless otherwise blocked by the theory of grammar, clitisation is available to every phrase, and thus constitutes a relevant proof for LD.



So far, I have tried to show the convenience of limiting the concept of LD to those instances parallel to TOP, the parameter [ $\pm$ pronominal copy] apparently constituting the only factor of differentiation between both linguistic phenomena. In doing so, I have rejected the inclusion of sentential modifiers in the general label LD, and have suggested that the obvious affinity between some junctives and LD constituents should be handled in terms of prototypes. This proposal, which will be modified later, in the first place, gets rid of the problems of coping with some non-integrable LD segments within a transformational framework, and secondly, goes back to the times of phrase-structure rules (Ross 1967, Bach 1974, similarly, Chomsky 1965), in which LD and TOP were given the same transformational treatment:

(34)	X	NP	Y
	1	2	3 $\rightarrow$ (optionally)
Topicalisation:			
2#[	1	0	3]
Left-dislocation:			
2#[	1	2	3]
		[+pro]	

where # stands for Chomsky-adjunction,  
where 2 is not a vocative.

I am not, however, supporting such a claim, since the perspective adopted here is couched in Government-Binding Theory (GB), according to which the only transformation is the well-known 'Move  $a_3$ ' constrained by general theoretical principles, to wit, subjacency, binding, and so on and so forth.

Neither should the reader believe that, in the light of the previous paragraphs, LD and TOP are almost identical. In Table 1 I put forward the behavioural differences between both phenomena identified in the literature, the justification of which falls beyond the scope of this paper (see, in particular, Rodman 1974 and Greenberg 1984):

*Table 1*

	LD	TOP
in questions	possible	impossible
in imperatives	possible	impossible
of NPs directly dominated by IP	possible	impossible
of complements of prepositions	possible	impossible
of quantified NPs	impossible (exceptions)	impossible
insertion of an introducer ( <i>as for...</i> )	possible	impossible
complex NP constraint	not aware	aware

	LD	TOP
gap	impossible	possible
subjacency	not aware	aware
<i>wh</i> -island constraint	not aware	aware
coordinate structure constraint	not aware	not aware(?)
sentential subject constraint	not aware	not aware
left-branch condition	not aware	aware
placement of interjections	after LD constituent	before TOP constituent

#### 4. Analyses of LD and TOP

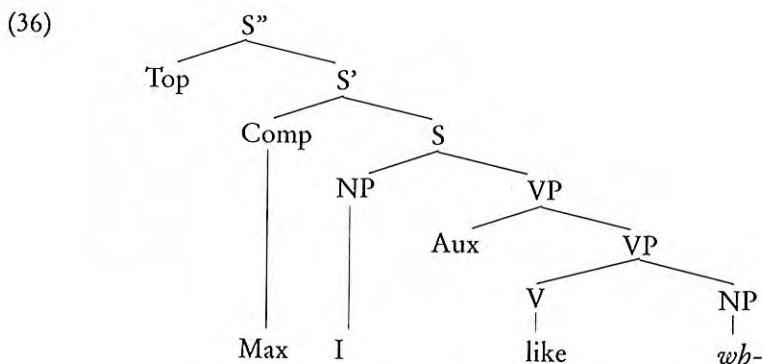
The differences outlined in the previous section led scholars to postulate independent sources for LD and TOP: base-generation in the case of the former, and a movement transformation in that of the latter.

The generation of LD constituents in the base, as proposed by scholars conversant with the transformational framework (e.g. Radford 1988, Haegeman 1994) poses no difficulties, the question of introducers evidently supporting such a possibility. The segment appearing in sentence-initial position is claimed to be adjoined to the sentence.

On the contrary, the sort of movement involved in TOP structures has aroused much controversy. In what follows, I shall restrict myself to discussing Chomsky's (1977) proposal, and how its weaknesses can be solved within the GB framework. Chomsky (1977) accommodates topicalised segments in the Comp node of *S'* in the phrase-structure rule in (35), whereas LD constituents are already in the base under Top:

- (35)  $S'' \rightarrow \text{Top } S'$   
 $(S' \rightarrow \text{Comp } S)$ ,

in which one must "take *S'* to be the maximal projection in that *S''*", i.e., *S* "would be an expansion of *S'*" (Chomsky 1981: 140, footnote 27). Example (36) sketches the process undergone by an underlying structure like *I like Max* in order to obtain the TOP *Max I like*:



In (36), *Max* is first generated under the NP of the VP, and *wh*-moves directly to Comp, leaving a trace coindexed to it.

Why is Top necessary? At first blush, one can easily observe that a node different from Comp is required in order to cope with examples like: (37), in which the Comp node is already filled by a *wh*-pronoun; (38), which shows the complementiser *that* in Comp position; and (39), in which either the topicalised phrase *under no circumstances* or the inverted auxiliary *would* already occupies the Comp node:

(37) He's a man to whom liberty we could never grant.<sup>8</sup>

(38) It's obvious that Mary, he can't stand.

(39) John swore that under no circumstances would he accept their offer.

The topicalised phrases in (37) and (38) appear in a position following Comp,<sup>9</sup> which is at odds with the phrase-marker suggested by Chomsky. Such a drawback led, for instance, Baltin (1982) to analyse Top as an adjunction to S:

(40) Move [X'] to S  
[-V]

GB theory has breathed new life to the problem, since the locations available as landing-sites for TOP are somewhat enriched. The Complementiser is claimed to be dominated by the Complementiser Phrase (CP), which has a Specifier. Constituents moved by *wh*-movement finally occupy the (Spec, CP) node just mentioned. Within this framework, TOP is seen as movement to (Spec, CP). Consequently, (37) could, in principle, be explained by ascribing *liberty* to the C node (formerly, Comp). However, (38) shows that this hunch does not render possible, since *that* is doubtlessly generated under C. What is more, if, following,

<sup>8</sup> Emonds (1976), in the light of examples like (i), postulates that TOP is only possible in root sentences: (i) \**Have I shown you the broom (that) these steps I used to sweep with?*. However, as reflected in example (37), TOP is compatible with, for instance, relative sentences. The problem with (i) is mostly dependent on stranding matters of the preposition *with*. In fact, the acceptability degree of (ii) is higher: (ii) *Have I shown you the broom with which these steps I used to sweep?*.

<sup>9</sup> McCawley (1988) points out that topicalised segments are perfectly acceptable after the complementiser *that*, but not after *whether* or a *wh*-expression, as respectively shown in (i) to (iii): (i) *I'm fairly sure that the skates John put in the closet*; (ii) ??*Fred asked whether the skates John put in the closet*; (iii) \**Fred asked where the skates John put*. In my view, both (i) and (ii) are possible utterances, generated by the adjunction of *the skates* to the sentence *John put *t* in the closet*. The problem with *wh*-relatives is not so clear-cut, since whereas (37) is perfectly grammatical, (iii) does not seem to be possible. In my opinion, the answer cannot be found on syntactic grounds. Since the topicalised constituent functions as a setting-definer with respect to the following sentence, or, in other words, the preposed material serves, in communicative terms, as the topic, to allow for two topics in these interrogative sentences does not make sense. The main (new) topic of (iii) is the interrogative *where*. In (37) above, *the man* is known information, so *liberty* is acceptable as the topic of the embedded clause.

The communicatively-biased criterion just put forward can well account for the different acceptability degree of (ii) with respect to (i). In sentences introduced by *whether*, the whole sentence, from the speaker's perspective, is informatively new. More plainly said, there is no place for a specific topic. On the contrary, *that*-clauses are more factual than hypothetical *whether*-clauses, so the preposing of topicalised material is allowed.

for instance, Chomsky (1986), we are to accept that subject-auxiliary inversion is a movement from I to C, then, in (39), *would* will fill the C node, as well as *that*. In consequence, another projection of CP is required in order to cope with example (39). Put more graphically, the structure would be:

(39') John swore [<sub>CP</sub> Spec [<sub>C</sub> that [<sub>CP</sub> [<sub>Spec</sub> under no circumstances] [<sub>C</sub> would<sub>C</sub> he...]]]]

Thus, one CP node is required in order to accommodate *that* in C, and another CP node is needed so that the TOP phrase can be properly placed under its Spec branch, the C node already filled by the inverted auxiliary. In fact, in order to deal with (37) or with interrogative examples like (41), taken from Ojea (1991), further projections of CP would be needed, since the (Spec, CP) node in these structures is filled by the *wh*-moved segment.

(41) Books, who need to read in the future?

This series of CP projections, known as 'CP interaction' (see Authier 1992) mitigates the problem of movement in TOP, though the complex machinery it requires remains a tantalising drawback.

As best I can determine, CP interaction is the only way of keeping track of TOP in a transformational way. I shall leave the question here since this paper is mostly concerned with the nature of LD- and TOP- nodes rather than with the specific machinery required by the theory of grammar in order to generate acceptable sentences with marked themes.

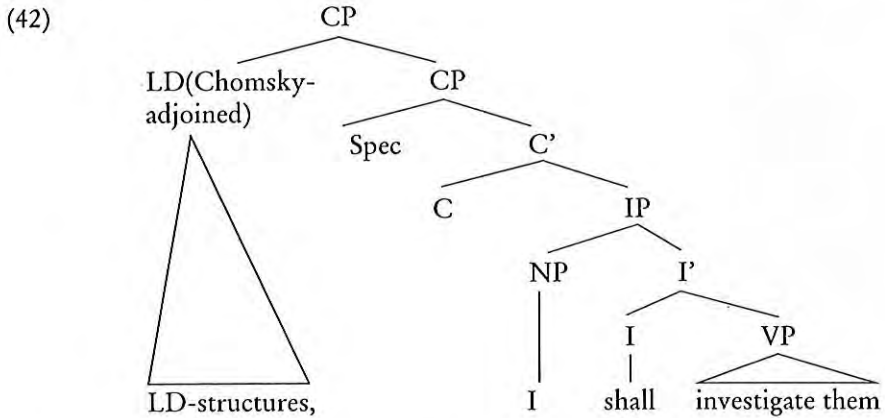
### 5. Generation of LD structures

In this section, I shall maintain that the node under which LD constituents generate must not be regarded as a murky node adjoined to whatever other category, but as a functional category itself. This interpretation is in keeping with the fact that, whereas some languages make wide use of LD in average speech, others are not eligible for LD. Thus, taking advantage of the categorial selection that, for instance, Chomsky's (1993) 'Minimalist Program' offers, the price of allowing for a, let us say, LD node in a particular group of languages is absolutely feasible (see, for instance, Marantz 1995: 364). Besides, as I shall try to demonstrate later, this functional category LD may prove useful in order to account for other linguistic phenomena.

In what follows I shall suggest that the semantic relation between the LD node of an utterance and the related segment in the ensuing sentence has a quasi-predicative nature. Besides, a semantic filter will be postulated which, first, will account for the pragmatic consequences brought about by the preposing of a given constituent, and, second, will integrate the 'semantic/ pragmatic compatibility' subtype of LD, defined by my condition (c) in (17) above.

5.1. *The semantic filter*

Let us now suppose that the basic tree-diagram of a given language which makes use of LD is as in (42):



The tree-diagram in (42) obviously belongs to a language in which the [ $\pm$ LD] parameter is activated. As stated above, this meets no further difficulties in the framework of Chomsky (1993). However, the production of sentences which make use of the LD functional category has to be regulated somehow. Let us assume that (at least) a semantic filter is responsible for the discrimination of LD-structures. Put differently, the aforementioned filter will predict that, for instance, (43) is impossible and, consequently, must be blocked:

(43) \*The sky, I must investigate LD-structures.

A first elaboration of the filter conditioning LD-structures could be (44), where  $\alpha$  contains the propositional meaning of the post-comma segment, “=” is the semantic relation of identity, and “ $\llbracket$ ” must be read as ‘semantically speaking.’ Y stands for the semantic content of the LD segment in initial position in PF, whereas X symbolises the clausal constituent which is related to Y.

(44)  $\alpha$ , where ‘X=Y’

Why should dislocated elements be interpreted in a predicative way? In the first place, the semantic relation of predication links two constituents, for example, subjects and predicates in equative sentences, or members of Small Clauses (SCs).<sup>10</sup> In the same vein, in (17), by conditions (a) and (b), I have defined LD as a device by

<sup>10</sup> In Small Clauses, “the entire complement is interpreted as a *situation* or *event*, that is, the second constituent is attributed to the NP, as if the NP were its ‘subject’” (Stowell 1978: 465). For the concept of Small Clause, see, for instance, Haegeman (1994: 123ff).

means of which a preposed constituent is connected to a grammatical copy occurring in the clause. Intuitively, those elements linked by LD could be associated by means of the bipartite nature of predicative structures. Secondly, as Aarts (1992) points out, a certain semantic relationship holds between the first and the second parts of an SC, to wit, predication.<sup>11</sup> In other words, an SC like the one italicised in (45) could be paraphrased as in (46):

(45) They appointed *her head of the Department*.

(46) [<sub>SC</sub> her head of the Department]  
'*her* [i.e. 'she'] is (going to be) *head of the Department*'

As for dislocated elements, a similar semantic link binds the constituent in sentence-initial position and its copy in the sentence. In fact, the connection is completely equative in the examples of the 'syntactic compatibility' subtype of LD, as (47) reflects:

(47) LD structures, they are SCs.

'*They = LD structures*'

With little additional machinery, I shall try to handle, by means of (44), those examples of LD not only of the syntactic subtype but of the pragmatic subtype as well. To that end, the relationship of predication, as depicted in (44), has to be widened in order to cover the operator  $\subset$  (inclusion) together with the = (equality) one. That granted, (44) could be reformulated as in (48):

(48) [<sub>LD-structure</sub> Y, sentence] =  $\alpha$ , where 'X $\subseteq$ Y'  
( $\alpha$ =propositional meaning of *sentence*; and X belongs to *sentence*)

Let us further accept that Y is the communicative topic of the sentence. In fact, this is in keeping with the general assumption that the initial position of an utterance is defined as topical by nature, which is not at odds with either generative, functional (systemic) studies. If we admit that only one constituent can be the topic of a sentence, then one will easily understand the appearance of exclusively Y in first position and of X in an orthodox sentential location other than sentence-initial. Put another way, the oddity of examples like (49) comes from the fact that, from a semantic point of view, *LD-structures* and *they* are connected by the relation discussed in (48).

(49) LD-structures, they prove difficult.

If *LD-structures* is interpreted as the topic of the utterance, its topical nature obviously percolates to its coindexed partner *they*. Since subject-position, that is, (Spec, P'), is, by definition, the unmarked location of topical elements (see Lakoff

<sup>11</sup> Cf. Matthews (1981), who analyses copulative sentences as non-verbal structures (our SCs) that are afterwards mapped onto verbal constructions by means of the copulative marker *be*.

1971), the use of an LD-configuration to mark the topicality of *LD structures* is not justified. In fact, a sentence such as (50) would fulfil the same communicative purposes (49) does:

(50) LD-structures prove difficult.

Example (49) would only be useful if, in a given context, *LD-structures* in (50) had to be understood contrastively (not topically).<sup>12</sup>

Given the semantic requirements posited above, the syntactic configuration of LD sentences can be understood as a formal way of marking the thematic organisation of the clause. In other words, LD would be seen as a device that marks as topical the constituent (preferably not in an initial position) which is semantically linked with that in first position in the way SC components are.

The topical marking carried out by LD-structures can be incorporated to (48), as in (51):

(51) *Thematic system of LD*: Interpret X in a topical way in a structure of the form of  $[_{CP} Y_{LD} [_{CP} \dots X \dots]]$ , where 'X  $\subseteq$  Y'

This thematic system of LD would characterise as topical the italicised constituents in (2), (9), (10), (12) and (14'), repeated here for convenience, on the basis that the semantic relation between them and the pre-comma material is sensible to the semantic filter in (48):

- (2) LD segments, I shall investigate *them*. [*them* = LD segments']
- (9) As for fronting devices, I have always liked *topicalisation and left-dislocation*. [*topicalisation and left-dislocation*  $\subset$  fronting devices']
- (10) LD constructions showing syntactic compatibility, *they* trigger interesting consequences. [*they* = LD constructions showing syntactic compatibility']
- (12) As for London, *Trafalgar Square* is nice. [*Trafalgar Square*  $\subset$  London']
- (14') My work, I'm crazy about *it*. [*it* = my work']

Thus, the theory of grammar is endowed with a theoretical mechanism which blocks every type of semantic/pragmatic unacceptability without further machinery. Condition (c) in (17) turns out to be a natural consequence of the semantic filter in (48).

<sup>12</sup> With regard to the examples in which the SC semantic relation holds between the dislocated segment and the syntactic subject of the sentence, I do not share Lakoff's (1971) intuitions when he claims that, out of the blue, the usual interpretation of *Mary, she talked to her* is (i) rather than (ii):

- (i)  $Mary_p, she_i$  talked to  $her_i$ .
- (ii)  $Mary_p, she_i$  talked to  $her_i$ .

Quite on the contrary, I contend that it is (ii) and not (i) that reflects the hearer/reader's understanding of the sentence on most occasions, as corroborated by some tests I have made with native speakers. In my opinion, the justification of the reading in (i) can be found in the tendency shown by pronouns towards the selection of near referents. In fact, the reading of *Mary declared that she had never talked to her* was that in (iii) in absolutely every consultation made to English speakers:

- (iii)  $Mary_i$  declared that  $she_i$  had never talked to  $her_i$ .

Examples (13) and (15), repeated below, are not permissible under (48), and thus are excluded from the class of LD-structures, as already suggested in §3:

(13) My work, I'm going crazy. [*I*  $\not\subset$  *my work*']

(15) As far as I'm concerned, London is nice. [*London*  $\not\subset$  (*as far as*) *I* ('*m concerned*')] ]

## 5.2. Further evidence

My proposal is able to cover resumptive pronouns in both substandard *wh*-clauses in English (52), sentential subjects (53), and introductory *its* depending on verbs like *deny* (54):

(52) The man who<sub>i</sub> John saw him<sub>i</sub>...

(53) He will come; that is clear.

(54) He denies it that John stole the money.

Besides, it supports Quirk et al.'s (1985) analysis of LD as a subtype of the appositive relationship, as well as Postal's (1994) judgements on TOP.

### (a) Resumptive pronouns

As remarked by Haegeman (1994: 409-10), in substandard English, *wh*-pronouns can 'undergo' LD. She correctly predicts that (52) cannot be regarded as a mere case of ill-formed relativisation in English, but as an instance of LD, since, as reflected in (55), that construction is not subject to subjacency:

(55) The man who<sub>i</sub> [<sub>IP#</sub>they think that if [<sub>IP#</sub>Mary marries him then everybody will be happy.

If my proposal is accepted, these examples can be treated as LD-structures, the base-generated (!) item *who* functioning as topical and *him* being its copy in the sentence, ruled by (51).

On the other hand, Koster (1978), in an article on extraposition, gives the residual example in (53) above, which my proposal can manage by assigning an LD-structure to the sentential subject in topic-position plus the sentence containing *that* in (53). In fact, if we are to admit Koster's thesis that these sentential subjects are not generated in subject position but outside S, and that then they move to a position previous to the empty subjects, we can apply our LD-analysis even to cases like (56):

(56) [<sub>LD</sub> That he will come] *e* is clear.



The LD-analysis, thus, proves a valid tool in order to account for overt resumptive pronouns and for Koster's 'satellite sentential subjects' with or without explicit subjects.

Postal extends the resumptive-pronoun question to TOP constructions, in order to account for the so-called parasitic-gaps ( $pg_i$ ), exemplified in (57):

(57) That <sub>$t_i$</sub>  he asserted  $t_i$  without verifying  $pg_i$ .

In his words,

[w]ithout going into great detail (...), considerable evidence can be found for both topicalization and 'object-raising' constructions, supporting the view that an invisible (resumptive) pronoun exists in their complements. (...) The view I advocate is that *NP and clausal topicalization constructions involve extraction both of the topic and of an invisible pronoun, which is a resumptive form associated with the first extraction.* (pp. 67-68) [my italics]

The sort of movement recognised by Postal in TOP structures affects both the fronted constituent and the pronoun with which it corefers. That granted, I partially make use of Postal's remarks, although I will allow movement only if the pronoun is covert. Were the pronominal copy overt, then the process involved would not be TOP but LD. With this, I do not deny that both the 'surface' fronted constituent and its copy are first generated in sentence-initial position.

#### (b) *Double-object constructions*

As for utterances like (54) above, with *it* as well as another major constituent with the function of object, I suggest that the pronoun plus the phrasal object should be treated as terms in an LD-relation. More graphically put, (58) would reflect the structure underlying (54):

(58) [<sub>LD</sub> That John stole the money], he denies it.

#### (c) *Quirk et al.'s (1985) appositive LDs*

I shall not insist on the fact that an immediate advantage of my LD-proposal is that it stands up pretty well with Quirk et al.'s analysis of the relationship between, for instance, *that John stole the money* and *it*, in (58) above, as appositive.

### 6. *Concluding remarks*

In this paper, I have demonstrated that the transformational analysis traditionally given to TOP is flawed if applied to LD. The assumption that LD also involves 'movement' to sentence-initial position does not stand up well under the

onslaught of technical tests such as questioning, imperatives, complex-NP constraint, subadjacency or the *wh*-island constraint, inter alia. As a consequence, a base-generation analysis is advocated in the literature for LD constructions.

Under my proposal, as commonly agreed in the relevant literature, LD constituents are generated in sentence-initial position under the node LD. In order to accommodate this extra node LD within the syntactic/semantic structure of the sentence, the theory of grammar requires some sort of link between the LD constituent and another one in the clause (its pronominal copy or semantically related material). To that end, a semantic filter governing such a link has been suggested, which has proved able to handle both the syntactic- and the semantic/pragmatic- types of LD. In support of such an analysis, a number of factors have been adduced, namely, the quasi-predicative relationship holding between both constituents (equational in the case of the 'syntactic compatibility' subtype of LD), the parallelism between dislocated segments and appositive structures, utterances with resumptive pronouns, double-object constructions, and TOP.

Finally, this proposal allows us to identify the pragmatic function of topic and the sentence-initial position. I will not attempt, however, the rather pointless business of surmising that, from ample historical and typological evidence already discussed in §2 in passing, every topical argument is generated as an SC in sentence-initial position, and that, in the course of its evolution to actual speech, it undergoes a 'de-marking'-like process, in the sense of Givón (1976). Were this view embraced, some insight into the nature of the relationship between communication and grammar would be gained.

### *Works cited*

- AARTS, Bas. *Small Clauses in English. The Nonverbal Types*. Berlin: Mouton, 1992.
- AUTHIER, J.-Marc. "Iterated CPs and Embedded Topicalization". *Linguistic Inquiry* 23 (1992): 329-336.
- BACH, Emmon. *Syntactic Theory*. New York: Holt, Reinhart and Winston, 1974.
- BALTIN, Mark R. "A Landing Site Theory of Movement Rules". *Linguistic Inquiry* 3 (1982): 1-38.
- BARBER, Charles. *Early Modern English*. London: Deutsch, 1976.
- BARCELONA SÁNCHEZ, Antonio. "El tópicos desgajado en inglés: Motivación pragmática". *Atlantis* 10 (1988): 9-20.
- BATES, Elizabeth. *Language and Context: The Acquisition of Pragmatics*. New York: Academic, 1976.
- CHAFE, Wallace L. "Givenness, Contrastiveness, Definiteness, Subjects, Topics, and Point of View". In Charles N. Li ed., 1976: 25-55.
- CHOMSKY, Noam. *Aspects of the Theory of Syntax*. Cambridge, MA: MIT, 1965.
- "On Wh-Movement". In Peter W. Culicover, Thomas Wasow, and Adrian Akmajian eds. *Formal Syntax*. New York: Academic, 1977: 71-132.
- *Lectures on Government and Binding: The Pissa Lectures*. Dordrecht: Foris, 1981.
- *Barriers*. Cambridge, MA: MIT, 1986.

- “A Minimalist Program for Linguistic Theory”. In Kenneth Hale, and Samuel Jay Keyser eds. *The View from Building 20*. Cambridge, MA: MIT, 1993: 1-52.
- CINQUE, Guglielmo. *Types of A'-Dependencies*. Cambridge, MA: MIT, 1990.
- DECLERCK, Renaat. *Studies on Copular Sentences, Cleft and Pseudo-Clefts*. Leuven: Foris, 1988.
- EMONDS, Joseph E. *A Transformational Approach to English Syntax: Root, Structure Preserving, and Local Transformations*. New York: Academic, 1976.
- FOLEY, William A., and Robert D. VAN VALIN. “Information Packaging in the Clause”. In Timothy Shopen ed. *Language Typology and Syntactic Description. Vol I: Clause Structure* Cambridge: CUP, 1985.
- GELUYKENS, Ronald. *From Discourse Process to Grammatical Construction. On Left-Dislocation in English*. Amsterdam: John Benjamins, 1992.
- GIVÓN, Talmy. “Topic, Pronoun, and Grammatical Agreement”. In Charles N. Li ed., 1976: 151-188.
- “Topic Continuity in Discourse. An Introduction”. In Talmy Givón ed. *Topic Continuity in Discourse: A Quantitative Cross-Language Study*. Amsterdam: John Benjamins, 1983.
- Greenberg, Gerald R. “Left Dislocation, Topicalization, and Interjections”. *Natural Language and Linguistic Theory* 2 (1984): 283-287.
- Gruber, Jeffrey S. “Topicalization in Child Language”. In David A. Reibel, and Sanford A. Schane eds. *Modern Studies in English*. New Jersey: Prentice-Hall, 1969.
- HAEGEMAN, Liliane. *Introduction to Government & Binding Theory*. Oxford: Blackwell, 1994 (2nd ed).
- HALLIDAY, Michael A.K. “Notes on Transitivity and Theme in English, Part 2”. *Journal of Linguistics* 3 (1967): 199-244.
- HANKAMER, Jorge. “On the Noncyclic Nature of WH Clefting”. *CLS* 10 (1974): 221-233.
- HERNANZ, M<sup>a</sup> L., and J. M<sup>a</sup> Brucart. *La Sintaxis. 1. Principios teóricos. La oración simple*. Barcelona: Crítica, 1987.
- KOSTER, Jan. “Why Subject Sentences don't Exist”. In Samuel J. Keyser ed. *Recent Transformational Studies in European Languages*. Cambridge, MA: MIT, 1978: 53-74.
- LAKOFF, George. “On Generative Semantics”. In Danny D. Steinberg, and Leon A. Jakobovits eds. *Semantics. An Interdisciplinary Reader in Philosophy, Linguistics and Psychology*. Cambridge: CUP, 1971: 231-296.
- LI, Charles N. ed. *Subject and Topic*. New York: Academic, 1976.
- MARANTZ, Alec. “The Minimalist Program”. In Gert Webelhuth ed. *Government and Binding Theory and the Minimalist Program*. Cambridge, MA: Blackwell, 1995: 349-382.
- MATTHEWS, Peter H. *Syntax*. Cambridge: CUP, 1981.
- MCCAWLEY, James D. *The Syntactic Phenomena of English*. Chicago: U of Chicago P, 1988.
- OCHS, Elinor-K., and Bambi B. Schieffelin. “Foregrounding Reference: a Reconsideration of Left Dislocation in Discourse”. *Proceeding of the 2nd. Annual Meeting of the Berkeley Linguistics Society*. Berkeley: U of California, 1976: 240-257.
- OJEA LÓPEZ, Ana Isabel. “Topicalización vs. dislocación: Algunas precisiones”. In Santiago González y Fernández-Corugedo, Juan E. Tazón Salces, María Socorro Suárez Lafuente, César Alas García, and Virginia Prieto López eds. *Studia Patriciae Shaw Oblata. Vol. II*. Oviedo: U de Oviedo, 1991: 155-169.
- POSTAL, PAUL M. “Parasitics and Pseudoparasitic Gaps”. *Linguistic Inquiry* 25 (1994): 63-117.
- PRINCE, Ellen F. “Topicalization, Focus-Movement and Yiddish Movement: A Pragmatic Differentiation”. *Proceedings of the Annual Meeting. Berkely L.S.* 7 (1981): 249-264.

- QUIRK, Randolph, Sidney GREENBAUM, Geoffrey LEECH, and Jan SVARTVIK. *A Comprehensive Grammar of the English Language*. London: Longman, 1985.
- RADFORD, Andrew. *Transformational Grammar. A First Course*. Cambridge: CUP, 1988.
- RODMAN, Robert. "On Left-Dislocation". *Papers in Linguistics* 7 (1974): 437-466.
- ROSS, John R. *Constraints on Variables in Syntax*. Indiana, Bloomington: Indiana U Linguistics Club, 1967.
- STOWELL, Tim. "What was There before There was There?". *CLS* 14 (1978): 458-471.