# cunyforum

## PAPERS IN LINGUISTICS

Proceedings of the Ninth Annual Meeting of
the North East Linguistic Society

Part 1

Ed Battistella, Editor

# A Reanalysis of Topicalization in Spanish[1]

Elizabeth A. Cowper
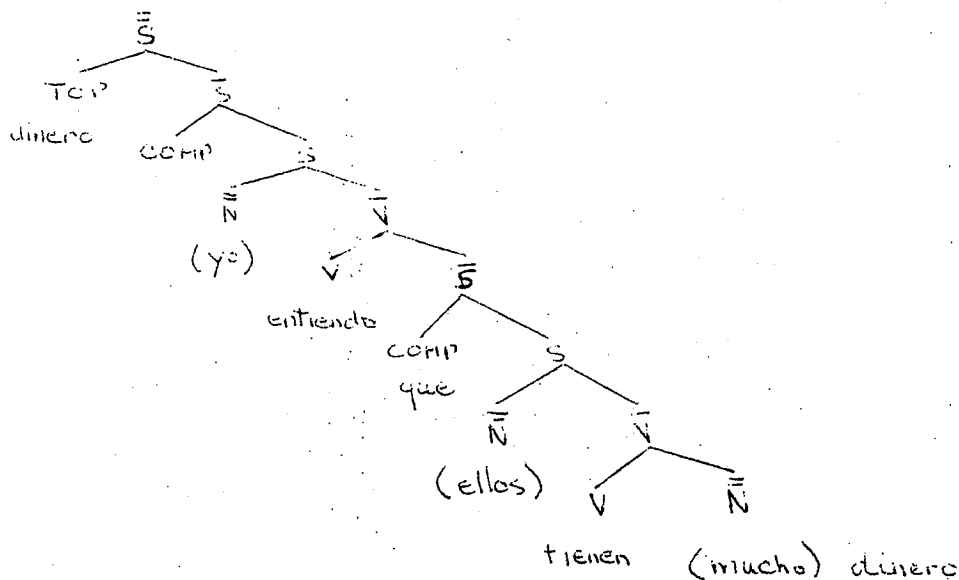University of Toronto

0.    The phenomena of topicalization and left-dislocation in Spanish[2] have been discussed at length in Rivero(1978) and Rivero(forthcoming). The present paper briefly summarizes these treatments, and shows that they are inadequate in one important respect. An alternative analysis for topicalization is proposed, and shown to be more adequate. Section 1 is a discussion of Rivero's analysis, section 2 demonstrates its inadequacy, and section 3 sets forth the alternative. Section 4 is a brief exploration of some of the problems with, and consequences of, the alternative proposed in section 3. The general framework of the Extended Standard Theory is assumed throughout.

1.0    Rivero(1978, forthcoming) has proposed the following analysis for topicalized sentences in Spanish:

1.1    Sentences like (1) and (2) have an underlying structure containing a sentence-initial TOP N'', and another N'' which is wholly or partially identical to the material in TOP.

    (1)   Dinero, entiendo que tienen.
          money understand+1s that have+3p

          Money, I understand that they have.

    (2)   Dinero, entiendo que tienen mucho.
          money understand+1s that have+3p much

          Money, I understand that they have a lot (of).

1.2    There is an optional rule which moves the lower N'' into COMP, and
thence from COMP to COMP.  Rivero calls this rule Topicalization, but to
avoid confusion with the name of the general phenomenon, I shall call it
T-movement.  T-movement shares some of the characteristics of WH-movement,
but, as Rivero shows, it cannot be identified with WH-movement.[3]  She
demonstrates that the Spanish complementizer has the structure shown in (3).

> (3)    [ QU $\Delta_t$  $\Delta_{wh}$ ]

Like WH-movement, T-movement observes subjacency and strict cyclicity.

1.3    An independently motivated transformation of deletion under identity
applies, to delete that part of the lower N'' which is identical to the
material in TOP.  Sentences like (2), (4) and (5) can thus be generated.

> (2)  Dinero, entiendo que tienen mucho.

> (4)  Dinero, entiendo que, mucho, tienen.

> (5)  Dinero, mucho, entiendo que tienen.

In (2), T-movement has not applied, and deletion has applied.  In (4),
T-movement has applied once, followed by deletion, and in (5), T-movement
has applied twice, followed by deletion.  The deletion rule is the same
rule as the one involved in the derivation of (6).

> (6)  Les dieron dinero a [$_{N''}$ los [$_{S'}$ que ya tenían $\emptyset$ ]]
>
>      to-them gave+3p money to them that already had+3p
>
>      They gave money to those who already had (money).

Since the deletion in (6) crosses two cyclic boundaries, Rivero claims that
the rule is unconstrained.

2.0    The above analysis correctly accounts for many cases of topicalization.
However, there is one class of sentences which it consistently fails to
handle.

2.1    Consider sentences (7) and (8):

> (7) *Dinero, entiendo [$_{N''}$ tu crítica [$_{S'}$ de que tienen $\emptyset$ ]]
>      money understand+1s your criticism of that have+3p

> (8)  Dinero, entiendo [$_{N''}$ tu crítica [$_{S'}$ de que tienen mucho $\emptyset$ ]]
>      money understand+12 your criticism of that have+3p much
>
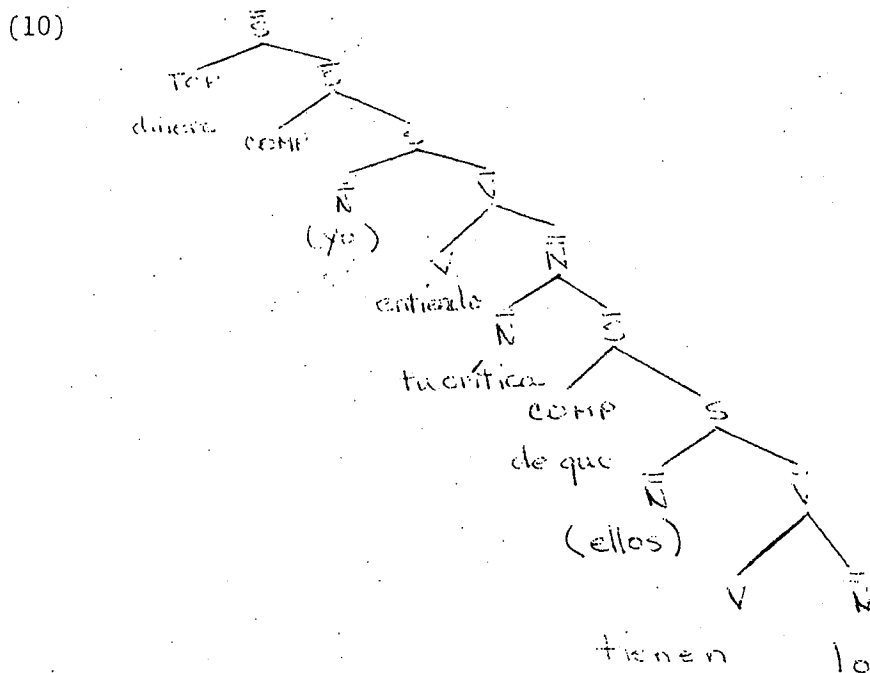>      Money, I understand your criticism that they have a lot (of it).

In both cases, there are two cyclic boundaries between TOP and the under-
lying position of the deleted noun.  The analysis will allow both sentences
to be generated, as follows:  In the case of (8), it is clear that T-movement
has not applied, since the quantifier is still in its underlying position.

The only rule that has applied, therefore, is the unconstrained rule of deletion under identity. With (7), it is unclear whether or not T-movement has applied. In any case, the only movement possible is into COMP in the embedded sentence; any further movement is blocked by the subjacency condition. However, nothing prevents (7) from being generated with a derivation exactly parallel to that of (8), involving no movement, but only deletion.

2.2    The crucial difference between (7) and (8) is that in (8), the lower N'' still contains lexical material, whereas in (7) it is empty. In this respect, (8) is more like a left-dislocated sentence such as (9), where the lower N'' contains a pronoun.

> (9)  El dinero, entiendo tu crítica de que lo tienen.
>       the money understand+1s your criticism of that it have+3p
>
>       The money, I understand your criticism that they have it.

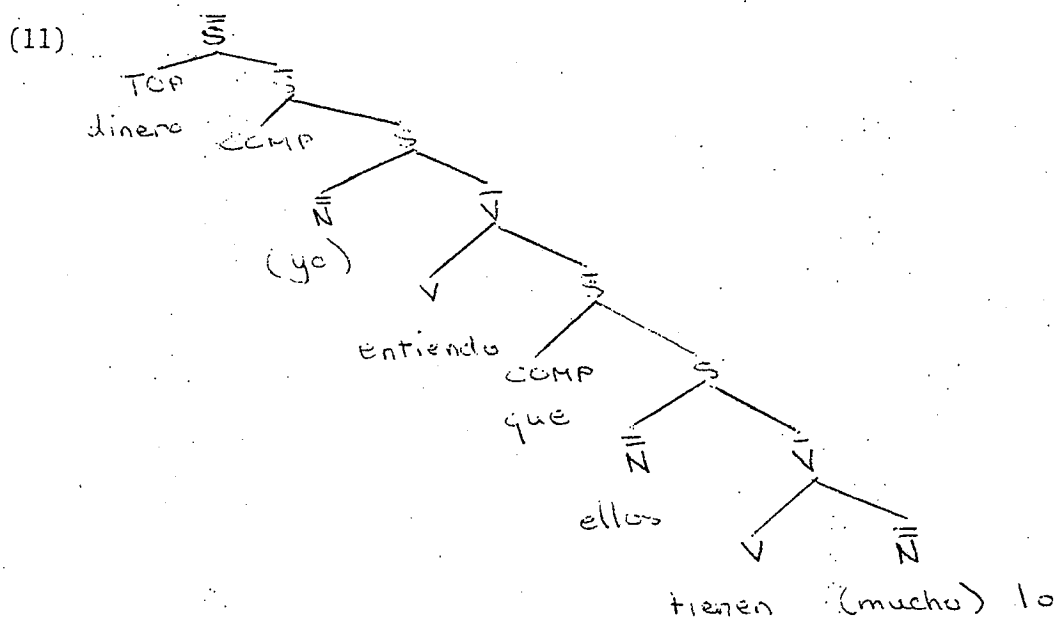Rivero derives sentences like (9) from structures like (10):

(10)



>    (lo cliticizes to the front of the tensed verb, as object pronouns
>    generally do in Spanish)

It would therefore seem that the distinction between left-dislocated and topicalized sentences is less clear-cut than one might think. The permissibility of complex NP structures between TOP and the lower N'' depends on the presence of lexical material in N'', rather than on whether the sentence is topicalized or left-dislocated.[4]

3.0    The following proposal will treat left-dislocation and topicalization as an essentially unified phenomenon, and will abandon the unconstrained rule of deletion under identity.

3.1    First of all, I propose that the underlying structure for sentences like (1) and (2) is as shown in (11):

(11)



This has two effects: First, it establishes an underlying structure for topicalized sentences which is essentially identical to that of left-dislocated sentences.  Second, it eliminates the need for the deletion rule to delete fully specified lexical material.

3.2    The derivation of (2) is as follows:  An obligatory deletion rule applies, to delete the pronominal head in the presence of mucho.[5]  This rule is local to N'', and is needed elsewhere in the language, as shown by (12) and (13).

    (12)    No tiene mucho, pero es felíz.
           not have+3s much but be+3s happy

      ·He doesn't have much, but he's happy.

242

(13)  Esta chica no quiere salir con feos.
      this girl not want+3s to-go-out with ugly+pl

      This girl doesn't want to go out with ugly (people)

Any N'' with a pronominal head undergoes this deletion rule if the N'' cont-
ains other lexical material.  Since this deletion rule is local to N'', it
is predicted that the structure intervening between the deletion site and
TOP (or any other constituent, for that matter) should not affect the
grammaticality of the sentence.  The derivation of (8) involves this same
local deletion rule, and (8) is therefore correctly predicted to be
grammatical.  Sentences (4) and (5) are derived by T-movement and pronominal
head deletion.  These two rules cannot be crucially ordered, since one moves
an entire N'', and the other applies strictly within N''.

3.3    Sentence (1) cannot be derived by pronominal head deletion, since N''
contains no lexical material other than the pronominal head.  Its derivation
is as follows:
       T-movement applies to move lo into COMP in the lower S', and then
again to move it into the higher COMP.  At this point, a deletion rule,
which I shall call T-deletion, applies to delete a pronoun in COMP if that
COMP is adjacent to TOP.  This deletion rule itself seems to have no true
independent motivation, but it is not unlike the rule which deletes WH-words
in Spanish relative clauses, illustrated in (14) and (15).

          (14)a.El hombre a quien dí el libro...
                the man to whom gave+1s the book

              The man to whom I gave the book...

             b.¿A quien diste el libro?
               to whom gave+2s the book

              To whom did you give the book?

          (15)a.El hombre que me dió el libro...
                the man that me gave+3s the book

               The man that gave me the book...

             b.¿Quien te dió el libro?
               who you gave+3s the book

              Who gave you the book?

It can easily be argued that que in (15) is the QU complementizer, and that
the WH-word has been deleted.

3.4    The analysis presented so far correctly predicts the non-occurrence
of sentences like (7).  The pronominal head deletion rule cannot apply,
since the N'' contains no other lexical material.  T-deletion cannot apply,
since lo cannot move up as far as the COMP node adjacent to TOP; it is
blocked by the subjacency condition.

243

4.0     Several questions are raised by the proposal presented in §3, and
these must be answered before the analysis can be accepted:
        What is the status of the unconstrained deletion rule? Sentences
like (6) must be generated somehow -- how can the process which generates
them be prevented from generating sentences like (7)?
        What does the grammar have to say about derivations where a pronoun
undergoes T-movement, but cannot move as far as the COMP adjacent to TOP,
and gets stranded in an intermediate complementizer?  (16) is an example
of such a case:

> (16) *Dinero, entiendo tu crítica de que, lo, dicen que tienen.
>       money understand+1s your criticism of that it say+3p that have+3p

How does the grammar handle the fact that 'left-dislocation' is
restricted to definite and generic noun phrases, and 'topicalization' is
restricted to indefinite noun phrases?

4.1.1 Concerning the unconstrained deletion rule, it should be pointed out
that it is not entirely unconstrained. First, oblique noun phrases, the
deletion of which would leave a stranded preposition, may not be deleted.

> (17) *Les dieron dinero a los que habían salido sin Ø
>       to-them gave+3p money to them that have+3p+pst gone-out without
>
> They gave money to those who had gone out without (money).

Second, there seem to be some constraints on the structure intervening
between the two identical N''s:

> (18) *Les dieron dinero a los que habían escrito libros sobre la
>       to-them gave+3p money to them that had+3p written books about the
>
> gente que necesita Ø
> people that need
>
> They gave money to those that had written books about the people
> who need (money)

We will not be concerned here with the exact statement of these constraints,
but will merely note that they exist. Third, since there is a productive
rule of subject pronoun drop in Spanish, the unconstrained deletion rule
need never be invoked to account for a missing subject.
        It would therefore seem that the only N''s affected by this deletion
rule are direct objects. Further, there is no reason to assume that the rule
deletes fully specified N''s. The rule could be stated at least as easily,
and perhaps more easily, if what was deleted was a pronoun.

4.1.2 A major distinction between sentences like (6) and those containing
an initial TOP is that an N'' in TOP is obligatorily coreferential with
another N'' in the sentence. Sentence (19), where the N'' in TOP has no
coreferent N'', is ungrammatical.

(19) *Dinero, creo que Rosita está enferma.
     money believe+1s that Rosita is+3s sick

There is no obligatory coreference in (6), and sentence (20) is therefore grammatical:

(20) Les dieron dinero a los que ya tenían muchos libros.
     to-them gave+3p money to them that already had+3p many books

     They gave money to those that already had many books.

One could therefore claim that there is a rule which deletes object pronouns if they are anaphoric and not in a relationship of obligatory coreference. This rule is subject to other conditions as well, but they will not concern us here.

4.2    Concerning the problem of sentences like (16), I must confess that I do not have a satisfying solution. I can only propose that sentences in which COMP dominates a -WH pronoun be filtered out at the level of surface structure.[6]

4.3    With regard to the definite-indefinite restriction, there are two possibilities. The facts are as shown in (21)-(26); (21) and (22) are from Rivero(forthcoming, p. 35).

(21) Libros, dicen que lee Ø
     books say+3p that reads+3s

     Books, they say that he reads.

(22) *Los libros, dicen que lee Ø
     the books say+3p that reads+3s

     The books, they say that he reads.

(23) Una manzana, dice que quiere comer.
     an apple says+3s that wants+3s to-eat

     An apple, he says that he wants to eat.

(24) *Una manzana, dice que quiere comerla.
     an apple says+3s that wants+3s to-eat-it

     An apple, he says that he wants to eat it.

(25) Dinero, entiendo que dicen que tienen.
     money understand+1s that say+3p that have+3p

     Money, I understand that they say that they have.

(26) Dinero, entiendo que dicen que lo tienen.
     money understand+1s that say+3p that it have+3p

     Money, I understand that they say that they have it. (GENERIC)

Since generically interpreted noun phrases pattern with definites with respect to this phenomenon, I will henceforth use the word definite to include both definite and generic.

4.3.1  The first possibility that comes to mind is to propose that there are two kinds of pronouns in Spanish: definite and indefinite. $\alpha$ definite N''s can only be in a relationship of obligatory coreference with $\alpha$ definite N''s. This restriction excludes (24), since la is a definite pronoun, and una manzana is an indefinite N''.
        T-movement could be restricted to move N'' only if N'' is indefinite. This would exclude (22), since (22) would involve the movement, and subsequent deletion, of a definite pronoun.  If the pronoun were indefinite, the sentence would be excluded by the above restriction on coreference. Restricting T-movement to indefinites provides an added bonus: it prevents sentences like (16) from being generated, and thus allows us to do away with the filter for COMP dominating -WH pronouns.
        The only problem with this alternative is that there are no surface indefinite pronouns in Spanish, and therefore the pronoun we are proposing is an abstract lexical item with no phonological shape.  Sentences whose underlying structures contain this pronoun, and whose derivations have not succeeded in deleting them, would have to be filtered out.
        This alternative accounts for all of the data, but it has three disadvantages:
- an abstract lexical item
- T-movement needs to refer to the definiteness of the moved N''.
- A filter is required to throw out sentences containing surface occurrences of the abstract lexical item.

4.3.2  A second way of accounting for the data is as follows:  First, all of the pronouns in question are definite.  There are no abstract pronouns. T-movement can therefore not be restricted to moving only indefinite N''s, since there are no indefinite pronouns in the language.  On the other had, T-deletion could be made sensitive to the definiteness of the N'' in TOP; it would apply only when that N'' is indefinite.  This would exclude (21); since los libros is definite, T-deletion would not apply.  A statement that $\alpha$ definite N''s can only be in a relationship of obligatory coreference with $\alpha$ definite N''s, which the other approach also requires, would exclude (24).  However, (16) now returns to haunt us.  We would still need a filter to block sentences in which COMP dominates a -WH pronoun.
        The second alternative, like the first, accounts for all of the data, but it also has disadvantages:
- T-deletion must refer to the definiteness of the N'' in TOP.
- a filter is required to block sentences where COMP dominates -WH pro.

4.3.3  Both alternatives require a statement about the definiteness of N''s involved in relationships of obligatory coreference.  Both alternatives have a transformation which refers to the definiteness of an N'': T-movement in the first alternative, and T-deletion in the second. However, in the first case, the N'' referred to must in any case be a term of the transformation. In the second case, the N'' in TOP could otherwise be left out of the structural index of the transformation; reference could be made merely to the presence of the TOP node.  This difference seems to argue in favor of alternative 1.
        Both alternatives require a filter.  Alternative 1 must propose an

abstract lexical item, while alternative 2 need not. All other things being
equal, this would argue for alternative 2. However, all other things are
not equal. We must therefore evaluate the relative nastiness of the T-dele-
tion rule in alternative 2 and the abstract pronoun in alternative 1.

Since the T-deletion rule violates the standard well-formedness
conditions on transformations, and since lexical items with no phonological
shape have been seen elsewhere, alternative 1 is to be preferred.

5. The analysis of topicalization, then, can be summarized as follows:
Topicalized N''s are base-generated in sentence-initial position. There is
an indefinite pronoun elsewhere in the sentence. This pronoun may be deleted
in place, if the N'' dominating it also dominates other lexical material.
T-movement can optionally move an indefinite N'' into COMP, and from COMP
to COMP, provided subjacency and strict cyclicity are observed. A rule
of T-deletion obligatorily deletes a pronoun in a COMP adjacent to TOP.
Two restrictions are needed, one to specify that $\alpha$ definite N''s can only
be in a relationship of obligatory coreference with $\alpha$ definite N''s, and
one to filter out sentences containing surface indefinite pronouns.

## Notes

1.  I am indebted to Juana Liceras and Carlos Yorio for their help as informants, and to Michael Szamosi and Aleks Steinbergs for comments and discussion on the preliminary version of this paper.

2.  The dialect of Spanish under discussion is Castilian Spanish. Latin American Spanish differs from Castilian in many ways, not the least of which is the behavior of topicalized sentences.

3.  Specifically, Rivero demonstrates that T-movement is not subject to the WH-Island constraint. It is sensitive to a T-Island constraint, which in turn does not affect WH-movement.

4.  The possibility that the ungrammaticality of sentences like (7) is related to perceptual constraints is not ruled out. I expect that the subjacency condition may in fact follow from a general theory of syntactic processing. The investigation of that question, however, is beyond the scope of the present paper.

5.  A similar process has been proposed by Ford and Singh (1978) for several languages, including Spanish.

6.  This problem will disappear as a result of the solution to the third problem.

## References

Ford, A. and R. Singh (1978). "On Non-Nominal Anaphora" Paper presented to the Canadian Linguistic Association, London, Ontario.

Rivero, Maria-Luisa (1978). "Topicalization in Spanish" Paper presented to the Canadian Linguistic Association, London, Ontario.

Rivero, Maria-Luisa (forthcoming) "On Left-dislocation and Topicalization in Spanish"