

North East Linguistics Society

Volume 11 *Proceedings of the Eleventh Annual Meeting of the North Eastern Linguistic Society*

Article 4

1981

On Extraction from Clitic Doubled Constructions

Hagit Borer

Follow this and additional works at: <https://scholarworks.umass.edu/nels>

 Part of the [Linguistics Commons](#)

Recommended Citation

Borer, Hagit (1981) "On Extraction from Clitic Doubled Constructions," *North East Linguistics Society*. Vol. 11 , Article 4.

Available at: <https://scholarworks.umass.edu/nels/vol11/iss1/4>

This Article is brought to you for free and open access by ScholarWorks@UMass Amherst. It has been accepted for inclusion in North East Linguistics Society by an authorized editor of ScholarWorks@UMass Amherst. For more information, please contact scholarworks@library.umass.edu.

ON EXTRACTION FROM CLITIC DOUBLED CONSTRUCTIONS

HAGIT BORER

MIT

0. Introduction

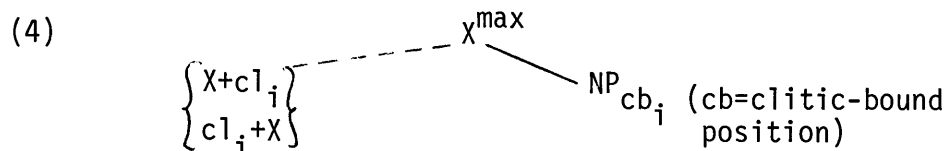
0.1 As a point of departure, consider the following sentences:

- (1) Lo_i vimos a Guille_i (River Plate Spanish (RPS), Jaeggli
him saw-we to Guille⁽¹⁹⁸⁰⁾
"we saw Guille")
- (2) L_i-am văzut pe Popescu_i (Rumanian, Steriade (1980))
him-have-I seen acc Popescu
"I have seen Popescu"
- (3) ze beit-o_i šel ha-'iš (Modern Hebrew (MH))
this house-his of the-man
"this is the man's house"

The sentences (1)-(3) demonstrate the phenomenon known as clitic doubling (henceforth CD). In particular, note that in each sentence an argument appears both as a clitic and as a full NP. In sentences (1),(2) the clitic is attached to the verb and in (3) the (genitive) clitic is attached to a noun.

Aoun (1979) and Jaeggli (1980), following an observation of Kavne, suggest that the structure of CD constructions is as in (4):

HAGIT BORER



It is argued that in (4) the clitic absorbs the Case features of the category X (or is itself the spell-out of the Case features). Jaeggli₁ further argues that the clitic absorbs the government properties of X. Crucially, the NP under X^{\max} (marked henceforth as NP_{cb}) is co-indexed with the clitic. Note that if one assumes the Case filter of Rouveret and Vergnaud (1980) as in (5):

- (5) *NP where NP is non-Case marked

then no lexical material can appear in NP_{cb} unless an independent device is found which can assign Case to it (recall that the Case features of X are absorbed by the clitic). Such a Case assigning device is the insertion of a dummy Case marker, which can be observed in (1)-(3): in RBS it is the preposition a, in Rumanian it is the specificity marker pe and in MH it is the genitive preposition şel. Failure to insert these dummy elements would lead to ungrammaticality²:

- (6) *Lo vimos Guille
 (7) *L-am văzut Popescu
 (8) *ze beit-o ha-'iš

In this paper, I will assume the essentials of Jaeggli's analysis, according to which the clitic in (4) absorbs both Case and (proper) government³. Most of the paper will be devoted to investigating the status of NP_{cb} in (4) and in particular, to an investigation of the relationship that hold between the clitic and NP_{cb} . It will be argued that although the clitic absorbs the (proper) government features of X in (4), NP_{cb} is nevertheless properly governed by the co-indexed clitic. The notion of proper government by clitics will be argued for in detail and it will be shown that with this notion it is possible to give a satisfactory account of some rather complicated extraction facts in Standard Arabic (SA), Rumanian and MH. In section 4 of this paper it will be suggested that in structures like (4) the status of the clitic is subject to parametric variations: whereas in MH, Rumanian and SA it is a non-argument position, in RPS it is an argument position. The properties of this suggested parameter will be explored and it will be suggested that it can account for the unavailability of NP_{cb} extraction in RPS.

0.2 The Empty Category Principle and NP_{cb} . Chomsky (forthcoming) argues that the following principle holds in universal grammar:

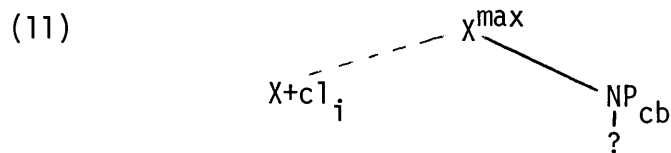
- (9) The Empty Category Principle (ECP):
 [e] must be properly governed

Following Kayne (1979) we will assume that ECP holds in Logical Form (LF). The definition of proper government is given in (10):

- (10) α properly governs β iff α governs β and:
 i. α is $\pm N$, $\pm V$; or
 ii. α is co-indexed with β

ON EXTRACTION FROM CLITIC-DOUBLED CONSTRUCTIONS

Now consider again the structure of (4) when NP_{cb} is not filled:



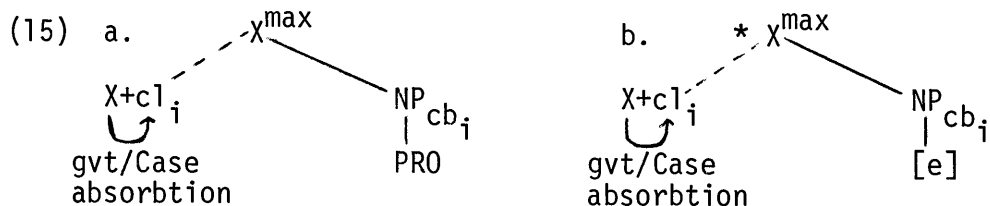
The structure in (11) is assumed to be the structure of (12)-(14):

- (12) Lo vimos (RPS)
him saw-we "we saw him"
- (13) L-am văzut (Rumanian)
him-have-I seen "I have seen him"
- (14) ze beit-o
this house-his ("this is his house")

A natural question arises as to the status of NP_{cb} : what does the question mark in (11) stand for? The following three hypotheses (at least) are logically possible:

Hypothesis A: NP_{cb} is an ungoverned position, hence only an ungoverned element can appear in it, namely PRO. Since the position is of course not properly governed as well, it follows that [e] cannot appear there, since it would violate the ECP. Thus extraction of NP_{cb} is impossible (this analysis is essentially due to Jaeggli).

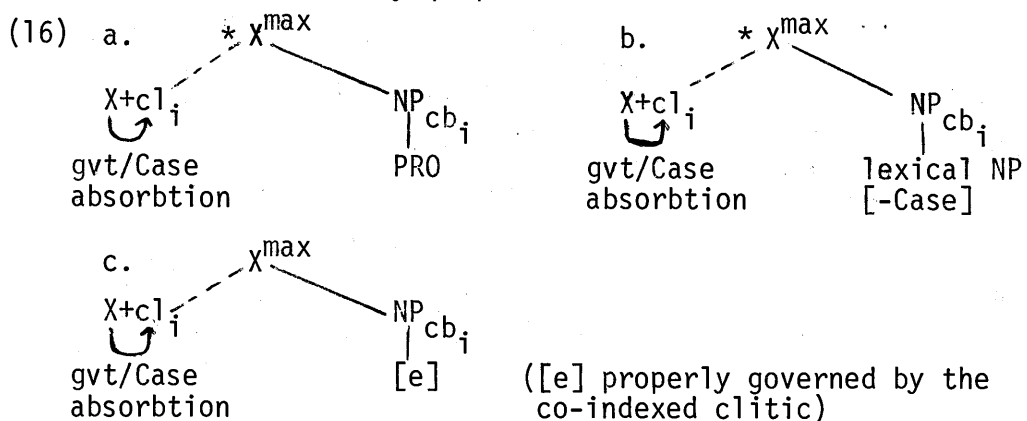
Hypothesis A is illustrated by (15):



In (15b), [e] is not properly governed hence the configuration would yield ungrammatical sentences.

Hypothesis B1: NP_{cb} is properly governed by the c-commanding, co-indexed clitic, hence PRO cannot appear there, lexical NP cannot appear there but [e] can. Extraction of NP_{cb} is possible⁴.

Hypothesis B1 is illustrated by (16):



HAGIT BORER

Hypothesis B2: The clitic absorbs only Case but not (proper) government. NP_{cb} is (properly) governed by X, hence PRO cannot appear there, lexical NP cannot appear there but [e] can. Extraction from NP_{cb} is possible⁵.

The predictions made by hypothesis B2 are identical to those made by B1 in most respects. Thus (16a-c) demonstrate the relevant configurations for hypothesis B2 as well.

Note that hypothesis A differs from B1 and B2 in one crucial, highly testable respect: whereas A predicts that extraction of NP_{cb} is impossible, both B1 and B2 predict that it is possible. In section 1 of this paper I will argue that in at least some languages (MH, Rumanian), extraction of NP_{cb} is obviously possible. Thus I will argue directly against A and indirectly for either B1 or B2. In section 2 I will show that when the relevant sentences can be constructed, it can be demonstrated that B1 rather than B2 is the correct hypothesis, thus arguing directly for B1 and indirectly against B2. (Note that by arguing directly for B1 I argue indirectly against all the other logically possible hypotheses which claim that NP_{cb} is properly governed by an element other than the co-indexed clitic).

1. Evidence for Extraction (A vs. B)

1.1 Rumanian. CD in Rumanian is subject to some semantic constraints. In particular, it only occurs when NP_{cb} is [+specific/definite] and [+human/pronominal]. This condition can be demonstrated in (17)-(18):

(17) L-am văzut pe Popescu
him-have-I seen acc Popescu

(18) Caut un bucatar
I-am-looking-for a cook

In (17), NP_{cb} is definite and human, thus satisfying the relevant condition. In (18), although NP_{cb} is human, it is not definite or specific hence there is no doubling.

As shown by Steriade (1980), the very same conditions hold when NP_{cb} is extracted. This is demonstrated by (19)-(20):

(19) Pe care_i credeai ca am văzut-o_i t_i?⁶
acc which-one_i thought-you that have-I seen it_i?

(20) Ce_i credeai ca am văzut t_i?
what thought-you that have-I seen?

In (19) the fronted WH is specific and arguably, like all WH elements, pronominal. In (20), on the other hand, the WH word, although pronominal, is not specific, hence CD is blocked⁷.

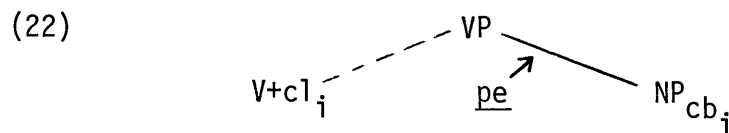
The situation in Rumanian can be summarized as follows:

(21) i(=17) cl_i V pe NP_{cb}_i
[+spec/def]
[+hum/pron]

ON EXTRACTION FROM CLITIC DOUBLED CONSTRUCTIONS

- ii(=19) pe WH_i.....cl_i V
[+spec./def]
[+hum/pron]
- iii(=18) V (pe) NP (elsewhere)
- iv(=20) (pe) WH_i.....V t_i(elsewhere)

Note that identical conditions apply to CD structures, whether or not extraction has applied and that in sentences such as (19) CD constructions such as (22) are the only possible source, thus indicating clearly that extraction of NP_{cb} in Rumanian is possible and that A makes false predictions for this language:

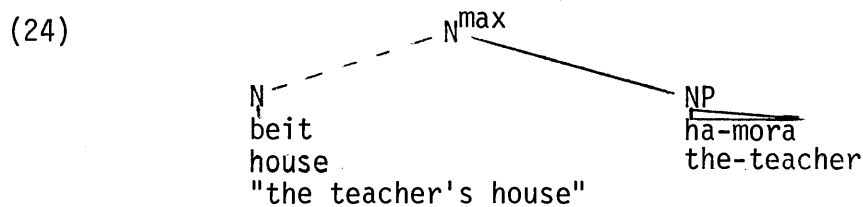


One could possibly argue that in fact, sentences such as (19) are no-movement cases and the clitic is actually a resumptive pronoun and the WH element is base-generated in place. However, as argued by Steriade (1980), conditions on movement cannot be violated in Rumanian, regardless of the existence of clitics. Thus (23), which is a Complex NP Constraint violation, is ungrammatical:

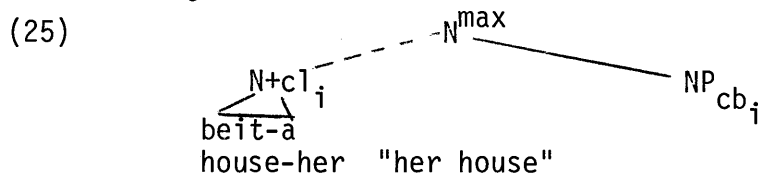
- (23) *Omul_j pe care_j o_j-conosc pe femeia_j care_j t_j l_j-a
 the-man acc whom her-know-I acc the-woman who t him-has
 întîlnit t_j a venit
 met t has come

Thus we conclude that the Rumanian extraction facts favor hypotheses B over hypothesis A.

1.2 Modern Hebrew (MH). CD in MH is attested only in one sort of construction: the construct state. The construct state, which roughly has the structure in (24), indicates genitival relations between the head N and the complement NP:

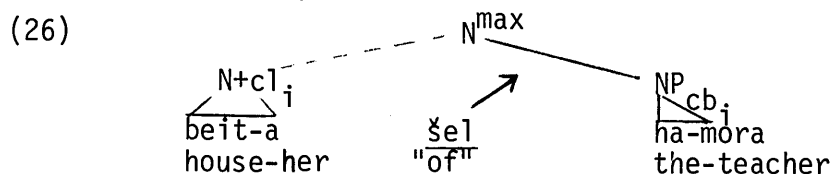


Nouns in MH (as in other Semitic languages) take clitics, which I will take to be the spell-out of genitive Case-marking features, otherwise assigned to the complement NP (this Case-marking actually surfaces morphologically in Standard Arabic). A noun+clitic combination will thus have the following structure:



HAGIT BORER

Structures such as (25) yield themselves to CD in the following way:

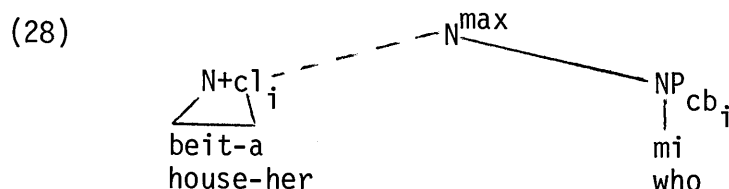


In (26), the genitive preposition *šel* has been inserted, thus enabling *ha-mora*, "the teacher" to receive Case (and see Borer (1980) where the properties of CD in MH are investigated in detail).

Now consider the free relative in (27):

- (27) *zo mi_j še-ra'iti 'et beit-a_j*
 this who that-saw-I acc house-her
 "this is the one whose house I saw"

Clearly, the co-occurrence of the WH word and the clitic could be explained only if we assumed that its source is the structure in (28):



Thus we conclude again that in MH extraction of NP_{cb} is possible⁸. Again, one could object that in (27) no movement has taken place and the clitic is, in fact, a resumptive pronoun. However, as in Rumanian, these constructions cannot violate constraints on movement:

- (29) **zo mi_j še-pagašti 'et ha-'iš_j še-ra'a 'et beit-a_j*
 this who that-met-I acc the-man that-saw acc house-her

This fact is particularly striking, since violations of constraints on movement are possible in MH in regular relative clauses, where arguably no movement has taken place. Thus the non-free relative corresponding to (29) is grammatical (and see Borer (1979) for discussion):

- (30) *zo ha-'iša_j še-pagašti 'et ha-'iš_j še-ra'a 'et beit-a_j*
 this the-woman that-met-I acc the-man that-saw acc house-her

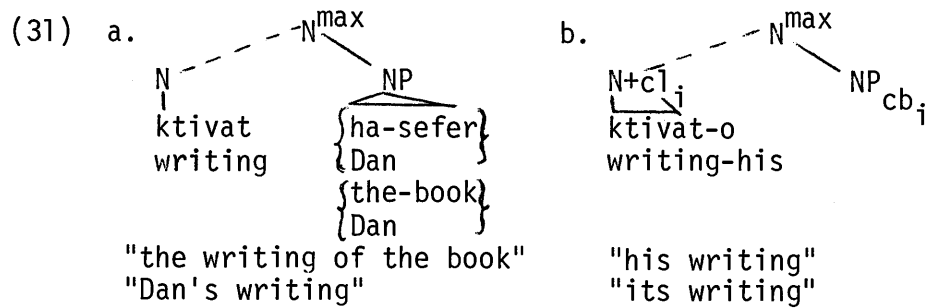
Since only in free relatives the constraints cannot be violated, we conclude that free relatives can be generated only by movement and that the WH word originated in the NP_{cb} position in (28). Thus MH as well supplies strong evidence against A, indirectly indicating that the correct hypothesis is B1 or B2.

2. Proper Government by Clitics (B1 vs. B2)

2.0 In section 1 it was shown that hypothesis A should be abandoned since it makes false predictions with respect to extraction in MH and Rumanian. In this section, I will argue that hypothesis B1 should be preferred over hypothesis B2. It will be shown that when two otherwise identical structures differ only with respect to the co-indexing of the clitic and NP_{cb} , extraction is possible only when co-indexing is present.

ON EXTRACTION FROM CLITIC DOUBLED CONSTRUCTIONS

2.1 Consider again the construct state in MH, which was illustrated in section 1.2 above. An interesting property of the construct state is that the NP_{cb} position is perfectly ambiguous: if the head noun can take both object and subject, NP_{cb} can be construed either as its subject or as its object:



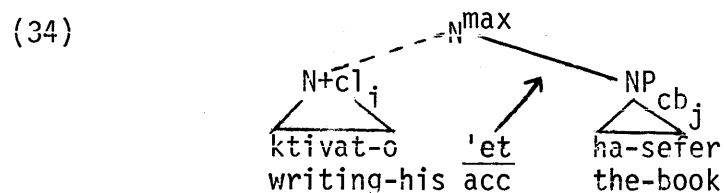
CD is possible regardless of the argument status of NP_{cb}. Hence both (32a) and (32b) are possible:

- (32a) ktivat-o_i šel Dan_i
writing-his of Dan
- (32b) ktivat-o_i šel ha-sefer_i
writing-it of the-book

Interestingly, when the clitic on the noun is understood to correspond to the subject, it is possible to generate the object in the position of NP_{cb}. Note that in this case the object argument will not have Case, since the genitive Case features are absorbed by the clitic. Hence a dummy accusative marker, 'et is inserted in order to assign Case to it⁹:

- (33) Dan biker 'et ktivat-o_i 'et ha-sefer_j
Dan criticized acc writing-his acc the-book

(33) has the structure (34):



The evidence that the accusative marker 'et is inserted in the underlined position in (33) is quite straight forward: the accusative marker 'et in MH is only present before definite direct objects. In other environments it is absent:

- (35) Dan 'axal 'et ha-banana
Dan ate acc the-banana
- (36) Dan 'axal (*'et) banana

In structures such as (34), the presence of 'et is obligatory. In effect, it results in the requirement that in (34), NP_{cb} should always be definite:

- (36) *Dan biker 'et ktivat-o_i ma'amar la-^Citon_j
Dan criticized acc writing-his article for-the-paper

HAGIT BORER

Note that the definiteness restriction on NP_{cb} is very strange and has no logical explanation other than the requirement that 'et be inserted for Case purposes and that 'et can only precede definite NP. (In English, for instance, the corresponding sentence with a non-definite object is grammatical: "John criticized his writing of an article for the paper").

We now have a minimal pair: (38) and (39) which differ only in one respect: while in (38) the clitic and NP_{cb} are co-indexed, in (39) they are not:

- (38) Dan biker 'et ktivat-o_i šel ha-sefer_i
Dan criticized acc writing-it of the-book
- (39) Dan biker 'et ktivat-o_i 'et ha-sefer_i
Dan criticized acc writing-his acc the-book

Whereas (38) has the structure in (26) above, (39) has the structure in (34). Do these sentences differ with respect to extraction? First note that if extraction has taken place, only the interpretation corresponding to (38) is possible¹⁰:

- (40) ma_j še-Dan biker 'et ktivat-o_i t_j
whatever_j that-Dan criticized acc writing-it_j
"whatever Dan criticized the writing of"
*"whatever Dan criticized his writing of"

However, an even clearer indication that extraction is possible in (26) but not in (34) is the contrast (41), (42). Recall that we assume that ECP holds in LF. Now consider a case in which a WH word in situ appears in NP_{cb} in sentences corresponding to (26) and (34)¹¹:

- (41) mi biker 'et ktivat-o_i šel ma/'eize sefer_i?
who criticized acc writing-it of what/which book
"who criticized the writing of what/which book?"
- (42) *mi biker 'et ktivat-o_i 'et ma/'eize sefer_i?
who criticized acc writing-his acc what/which book
"who criticized his writing of what/which book?"

If we assume hypothesis B1 above, we have an immediate explanation for the contrast of (41) and (42). Assuming that WH words in situ are subject to the rule of Quantifier Raising (in the sense of May (1977)) they leave behind an empty category that has to be properly governed. Consider now the LF representation for (41) and (42) (irrelevant details omitted):

- (43) i(=41) (for which x_i , x_i a thing/a book)...N+cl_i x_i
(proper government by clitic)
- ii(=42) (for which x_i , x_i a thing/a book)...N+cl_j x_i
(no proper government by clitic)

Assuming proper government by clitics, the nature of the violation in (43ii) is immediately clear: following the movement in LF of the WH in situ, the empty category x_i is not properly governed, since the preceding clitic is not co-indexed with it. In (43i), on the other hand,

ON EXTRACTION FROM CLITIC DOUBLED CONSTRUCTIONS

the preceding clitic is co-indexed with the empty category x_i , hence it can properly govern it and the sentence is not a violation of ECP.

To sum up our conclusions so far, we have determined on the basis of extraction facts in MH and Rumanian, that NP_{cb} in (4) above is properly governed, in spite of the fact that the clitic attached to X absorbs both Case and (proper) government. We have then determined on the basis of LF movement in MH that it is the co-indexed clitic which properly governs NP_{cb} , rather than any other element. Let us now consider the interaction of these conclusions in the analysis of Standard Arabic (SA) complementizer 'anna.

3. SA Complementizer 'anna

SA is essentially a VSO language. However, the complementizer 'anna (roughly 'that') triggers some interesting changes in word order. Thus 'anna appears in the following configurations (Aoun, 1980):

Configuration A: 'anna NP V NP
 [+accusative]
 (where the fronted NP is direct object or subject).

(44) 'aʕunnu 'anna lwalada_j 'akalat-hu_j ddifd^Catu (DO fronting)
 think-I that the-boy-acc ate-him the-frog-nom¹²

(45) 'aʕunnu 'anna ddifd^Cata 'akalat lwalada (subject fronting)
 think-I that the-frog-acc ate the-boy-acc

Note that the NP immediately following 'anna always receives accusative Case marking, regardless of the initial extraction site. If no fronting of an NP has taken place, 'anna is obligatorily followed by a clitic which is 3rd person singular masculine and which is arbitrary in reference. This is demonstrated by configuration B and (46):

Configuration B: 'anna+cl_i V NP_j NP_k

(46) 'aʕunnu 'anna=hu_i 'akalat ddifda^Catu_j lwalada_k
 think-I that-it_i ate the-frog-nom_j the-boy-acc_k

Let us assume that 'anna has roughly the following structure:

(47) a. b.

I will further assume, following Aoun (1980), that 'anna, as argued for prepositions by Kayne (1979) is not a proper governor. It follows immediately that, when 'anna is followed only by an empty category, as in (48), it is ungrammatical, as is demonstrated in (49):

(48)

HAGIT BORER

- (49) *'a^xunnu 'anna 'akalat ddifda^catu lwalada
 think-I that ate the-frog-nom the-boy-acc

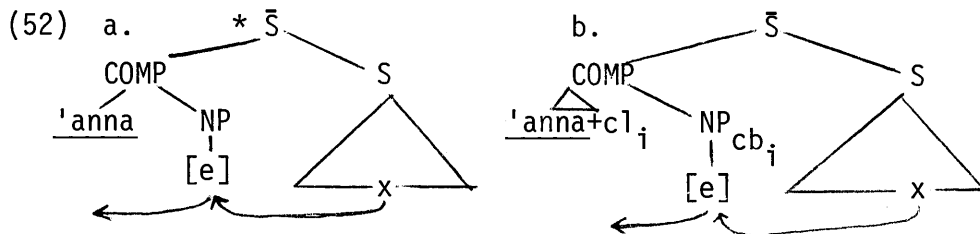
However, both a and b in (47) do not violate the ECP: in (47a) (and the corresponding sentences (44) and (45)) a full NP was preposed to the NP_{CB} position, receiving Case from 'anna and thus avoiding an ECP violation. In (47b) the co-indexed clitic properly governs the empty category, thus enabling a satisfaction of ECP.

When extraction takes place from 'anna clauses, an interesting change takes place: the clitic on 'anna which is otherwise arbitrary in reference can agree with the extracted element in gender number and person and in this case it is understood to be co-referential with it:

- (50) 'ayya: difda^catin ta unnu 'anna-ha. 'akalat lwalada?
 which frog-fem think-you that-her ate-fem the-boy?
 "which frog, do you think, ate the boy?"

(the clitic *ha* is feminine singular and thus agrees with the features of the antecedent *difda* at, "frog")

The conclusion that clitics properly govern a co-indexed empty category can explain these facts in a very simple way. Consider the structures in (52):



Clearly, only the extraction in (52b) is grammatical, since only in this case the empty category is properly governed by the co-indexed clitic. In (52a) such a co-indexed clitic is not present. A co-indexed clitic, however, has to agree in features with the element which it properly governs and note, that the NP_{CB} in (52b) now contains the trace of the extracted element. It follows that only when the clitic shares the features of the extracted element can it properly govern the NP_{CB} position in (52b), which now contains the trace of that extracted element. Thus only when the clitic agrees with the features of the moved phrase can it properly govern and thus avoid ECP violations.¹³

Thus the complementizer 'anna in SA further supplies evidence for proper government by clitics. Note that an alternative explanation of the requirement that the clitic agree with an antecedent only when extraction has taken place would have to resort to a process of shadow pronoun copying or some similar device. The assumption that clitics can be proper governors thus enables us to account for the data without such a mechanism. (For a solution involving a phonological spell-out of the trace as well as some discussion of the analysis above see Aoun (1980)).

ON EXTRACTION FROM CLITIC DOUBLED CONSTRUCTIONS

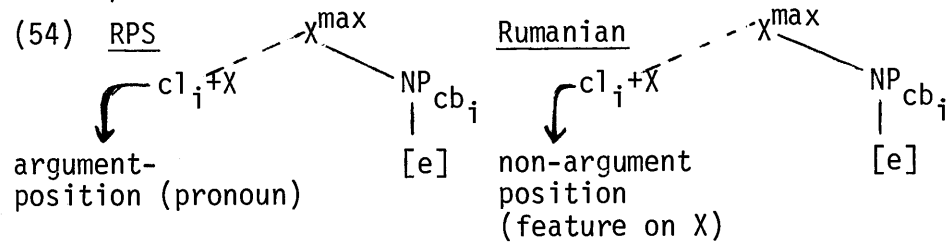
4. What Goes on in River Plate Spanish (RPS) (a possible parameter)

Jaeggli (1980) argues that in RPS extraction from clitic-doubled constructions in direct object configurations is impossible¹⁴:

- (53) *A quien lo viste?
 who him saw-you?

Clearly, RPS differs from Rumanian and MH in a crucial respect: extraction from NP_{cb} is impossible. Thus RPS seems to be a counter-example to our analysis of proper government by clitics.

Let us assume that RPS differs from Rumanian (with which it is identical in most other relevant respects) precisely in one way: while in RPS the clitic in CD constructions is interpreted as occupying an argument position, in Rumanian, MH and SA it is interpreted merely as a feature on the head of its phrase. Argument position (following Chomsky, class lectures Fall 1980) is essentially a position in which a thematic role can be assigned. In essence, what the difference between Rumanian and RPS implies is that in some sense the clitic in RPS is a pronoun whereas in Rumanian, for instance, it is not. Let us now compare the post extraction structures in the two languages:



Recall now that the binding conditions require (among other things) the condition in (55):

- (55) variables must be argument free

Further note that since the clitic in RPS is an argument, only in that language (but not in Rumanian) the relation between the clitic and [e] is that of argument binding. We are now able to explain why WH extraction of NP_{cb} is impossible in RPS: such extraction would leave a variable under NP_{cb} and this variable would be argument-bound by the co-indexed clitic. Thus such movement is impossible. Note that we still assume that the NP_{cb} position is properly governed by the co-indexed clitic in RPS as well as in Rumanian. It follows that PRO cannot appear there. Rather, I will assume that [e] which appears in the RPS structure in (54) is simply a base-generated [e] which cannot be interpreted as a variable. Note that the extraction of NP_{cb} in Rumanian would not yield violations of (55), since the clitic is a feature rather than a pronominal element.

Our suggestion that languages can parametrize with respect to the classification of various nominal elements does not involve an increase in the power of the grammar: on the contrary. Obviously a child has to learn the classification of nominal elements as they enter his/her lexicon. Thus it is not costly to assume that the child learns to interpret the clitic in Rumanian as a non-argument

HAGIT BORER

and the clitic in RPS as an argument. In fact, in Rumanian an immediate and direct evidence is given to the child that the clitic cannot be an argument, since clitic doubling in extraction constructions is obligatory, as it is in simple sentences.

The availability of direct positive evidence with respect to the nature of the clitic in Rumanian vs. the lack of positive evidence from RPS to the contrary (presumably, a child does not hear starred sentences) may well indicate that the situation in RPS is the least marked case.

Interestingly, CD is obligatory in Rumanian, but not in RPS. It may very well be that alongside the introduction of CD as an obligatory rule (which is a relatively recent phenomenon; see Steriade (1980) for discussion of earlier stages of Rumanian), the clitic, always present, was indeed interpreted as a feature on the head (in this case, always the verb). In RPS, where CD is optional, the initial status of the clitic as a pronoun is still retained, since its occasional absence makes it impossible to interpret it as a mere feature on the verb¹⁵.

5. Summary

It has been shown in this paper that by utilizing the notion of clitics as proper governors it is possible to account for extraction facts in Rumanian, MH and SA. It has been further suggested that a parameter of core grammar might involve the classification of different elements lexically as pronominal or non-pronominal. Such classification is independently necessary and thus does not increase the power of the grammar. The argument in this paper as a whole also serves to reinforce assumptions made by the general model. In particular, it points out the importance of the Empty Category Principle and the utilization of proper government and supports the conclusion that ECP holds in LF. It further supplies evidence in favor of the view that clitics, indeed, should be viewed as a spell-out of Case and (proper) government features.

Footnotes

*I am indebted to J. Aoun, N. Chomsky, J. Higginbotham, R. May, D. Pesetsky, H. van Riemsdijk, B. Schein, D. Steriade and T. Stowell for very fruitful discussions of material in this paper.

¹The notion of government absorption is somewhat strange: one might plausibly argue that government is a structural property which cannot be spelled-out or absorbed. I refer the reader to Jaeggli (1980) where the notion of S-Government is developed and it is suggested that this is the relevant notion for determining proper government. For our purposes, it would suffice to assume that only proper government is absorbed and that proper government (unlike government) is a lexical property. Thus henceforth the term (proper) government will be used whenever the discussion touches on government absorption.

²In Rumanian and RPS, in fact, the elements a and pe independently exist in the grammar as specificity markers. Thus they would appear

ON EXTRACTION FROM CLITIC DOUBLED CONSTRUCTIONS

before the direct object even if doubling did not occur. However, it is possible to argue that it is precisely the availability of these markers and their ability to assign Case that makes doubling possible, even if they are actually base-generated rather than inserted. In MH, on the other hand, $\bar{\xi}el$ is clearly inserted, since it would not precede the complement of N in sentences such as (3) unless doubling has taken place. The difference between $\bar{\xi}el$ on one hand and pe in particular on the other hand can actually account for the "frontability" of the latter vs. the impossibility of fronting $\bar{\xi}el$ and see fn. 8 below for some further discussion.

³For the meaning of (proper) government see fn 1 above.

⁴Note that, in order to assume B1, we must crucially assume that the combination X+c1 does not branch. If it branched, the clitic could not c-command the co-indexed NP_{CB} and hence could not govern (or properly govern) it (and see also fn. 15 for some further discussion).

⁵In fact, hypothesis B2 can be argued against directly on the grounds that in MH, for instance, where clitic doubling occurs inside NP's, N is not a proper governor and hence "noun stranding" is ungrammatical (and see Borer (1980) for a detailed discussion). Hypothesis B2, however, serves as an "inclusive straw man" and can be understood to refer to any proposal under which an element other than the clitic properly governs NP.

⁶Note that in (19) the clitic appears post-verbally, whereas in (17), for instance, it precedes the verb. This is determined solely on phonological grounds: if the clitic starts with a vowel it will cliticize at the end of the verbal complex rather than attach to a vowel-initial auxiliary.

⁷The facts of Rumanian are actually more complicated, but in ways which are not necessarily relevant to the discussion in this paper. Thus when nominal pied piping which is specific but not [+hum/pron] takes place, CD is obligatory, although in simple sentences (where extraction has not taken place) a clitic does not appear preceding such nominal complexes. Steriade (1980) argues that CD in such constructions can be accounted for if we assume shadow pronoun copying (which never surfaces) and which copies the specificity feature of the antecedent WH onto the pronoun. I will diverge from this analysis only by claiming that a variable can satisfy the [+pron] feature without any copying rule. However, for the purposes of our argument this is not relevant: the relevant point is that in Rumanian extraction from CD constructions is clearly possible if the semantic constraints are met.

⁸As in the case of Rumanian, the extraction facts in MH are considerably more complicated, but in ways which are irrelevant to this paper. Thus extraction from CD constructions is only possible in free relatives (and not, say, in questions) due to interaction with the Case filter. Note that the extraction in MH would be from a non-Case position (and crucially, assuming that \bar{S} rather than S is a bounding node). We expect this extraction to not violate the Case filter only when the antecedent WH receives Case in COMP. Such Case assignment

HAGIT BORER

in COMP is only available to free relatives, thus making extraction from CD constructions in questions ungrammatical due to the Case filter. Note that in this case the variable trace will not be Case marked (see Borer (1980) for the discussion of the theoretical consequences of this fact). In Rumanian, due to the fact that the accusative Case marker pe is fronted along with the WH element, the problem of assigning Case to the antecedent WH element never arises. The fact that pe is fronted along with the WH element but șel cannot be fronted supplies further evidence that whereas pe is base-generated as a specificity marker, șel is inserted later, after move α (in the sense of Chomsky (1975)) has applied. Thus if WH element is fronted from a CD construction in MH, the environment for șel insertion is destroyed and the WH element cannot receive Case in this fashion (and see fn. 2)

⁹An interesting question arises with respect to these constructions in which both the subject and the object can appear: (i), in which the clitic refers to the object and the NP_{CB} to the subject, is ungrammatical:

- (i) *ktivat-oj 'et/șel/∅ Danj
writing-it Dan
"it's writing by Dan"

This could be explained by the fact that there is no dummy Case assigner which could be inserted preceding the subject. A "by" phrase can appear following "its writing" as in (ii):

- (ii) ktivat-oj^Cal yedey Danj
by "its writing by Dan"

but there are good arguments that in (ii) the "by" phrase is a full PP, and which are outside the scope of this paper. Another interesting question concerns the fashion in which the grammar "knows" whether to insert șel or 'et. D. Pesetsky (personal communication) pointed out to me that in some sense, șel can be perceived of as a proximate marker whereas 'et is an obviative one. I will not discuss these issues in greater detail due to space limitations.

¹⁰The fronting of 'et along with the WH element in (40) would not improve matters. This is expected since the fronting of 'et would still leave an empty category which needs to be properly governed.

¹¹One could argue that the ungrammaticality of (42) is due to some version of the Specified Subject Condition (first suggested in Chomsky (1973)). However, the claim is problematic. First it is not clear in what sense the MH sentence (42) has a subject in a sense which (32a) doesn't. Yet, in the sentence parallel to (32a), a WH word in situ is possible:

- (i) Mi biker 'et ktivat-oj șel mij
who criticized his writing of whom

Second, note that the English sentence corresponding to (42) is perfectly grammatical:

- (ii) Who criticized his writing of which book?

thus voiding the claim that the SSC can rule these sentences out.

It would be plausible to assume that in (ii) a violation of SSC is avoided by pied-piping in Logical Form of the entire phrase rather than the WH word in situ by itself.

ON EXTRACTION FROM CLITIC DOUBLED CONSTRUCTIONS

¹²Note that following fronting from direct object position a clitic is attached to the verb, thus giving the sentence a dislocated meaning. This fact, which can be shown to obey most of the constraints presented in this paper, will not be discussed here.

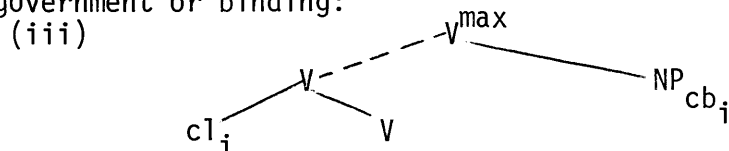
¹³A tacit assumption behind this statement is that clitic spell-out is random with respect to features such as gender, number and person. A generation of the wrong set of features is ruled out in such a system by the mismatch in features of two co-indexed elements. The spell-out of features as singular 3rd person masculine in SA, on the other hand, is taken to be the unmarked, non-referential case. Thus when the empty category in the complement of 'anna does not refer to anything, these features are generated and only when these features are generated the clitic is allowed to have arbitrary reference. Any other set of features leads to referentiality and thus requires that the clitic be co-indexed with an argument.

¹⁴CD in RPS and in Rumanian in non-direct objects is not discussed in this paper. It is interesting to observe, however, that in RPS, extraction from CD constructions other than direct object ones is possible.

¹⁵Barry Schein (personal communication) has pointed out to me that in fact, the imposing of the binding conditions on structures such as (54) in RPS would also entail that full referential NP's cannot appear in the NP_{cb} position, since the binding conditions also require that referential NP's be argument-free. There are two possible solutions to this problem: first, it is quite possible that the existence of the marker a functions precisely to block argument-binding when it is present. This would, perhaps, fall together with the intuition that a, as şel in MH, is a proximate marker (see fn. 9 above). Since a does not appear preceding variables, it could not block argument-binding in a similar fashion. Thus argument-binding would be possible in (i) but not in (ii):

- (i) Lo_i vimos [e]_i
 (ii) Lo_i vimos a Guille_i

Another possibility would be to claim that due to the argument status of the clitic in RPS the structure cl_i+X branches in RPS but not in Rumanian. A branching structure such as in (iii) would thus destroy the c-command relations between the clitic and NP_{cb} and would not enable government or binding:



Note that if this were the case, hypothesis A would have to be true for RPS, since NP_{cb} would be ungoverned and only PRO could appear there. However, note that a serious drawback to this analysis is the definition of PRO in Chomsky (forthcoming). According to this definition the difference between PRO and [e] ("NP trace") is that the former has a thematic role separate from that of its antecedent whereas the latter shares a thematic role with its antecedent. Further note,

HAGIT BORER

that a branching structure would empty a "spelling-out of features" analysis from its content. On these grounds it is obviously preferable to conclude that the branching proposal should be rejected and that one should adopt the assumption that a blocks argument-binding in sentences such as (ii) above.

REFERENCES

- Aoun J. (1979), "On Clitic Placement". Unpublished MIT ms.
 Aoun J. (1980), "ECP, Move and Subjacency", to appear in Linguistic Inquiry, Cambridge.
 Borer, H. (1979), "Restrictive Relatives in Modern Hebrew", Unpublished MIT generals paper
 Borer, H. (1980), "On the Definition of Variables" to appear in Journal of Linguistic Research, University of Indiana Linguistic Club, Bloomington.
 Chomsky, N. (1973), "Conditions on Transformations", in S. Anderson and P. Kiparsky (eds.) Festschrift for Morris Halle.
 Chomsky, N., (1975), "Conditions on Rules of Grammar", in Linguistic Analysis Vol. 2, No. 4.
 Chomsky, N. (forthcoming), Lectures on Binding and Government, Foris Publications, Dorcrecht
 Jaeggli, O. (1980), On Phonologically-null Elements in Syntax, unpublished dissertation, MIT, Cambridge
 Kayne, R. (1979), "ECP Extensions", to appear in Linguistic Inquiry May , R.C. (1977), The Grammar of Quantification, unpublished dissertation, MIT, Cambridge
 Rouveret A. and J.R. Vergnaud, "Specifying Reference to Subjects: French Causatives and Conditions on Representations", Linguistic Inquiry, Vol. 11, No. 1
 Steriade, D. (1980), "Clitic Doubling in the Rumanian WH Constructions and the Analysis of Topicalization", in A. Ojeda (ed.) Proceedings from the 16th Meeting of the Chicago Linguistic Society, Chicago 1980.