

A NOTE ON CATALAN CLITICS

Anna Gavarró

Universiteit van Amsterdam

The purpose of this paper is to suggest an interpretation of the colloquial clitic system of Barcelona Catalan (and other dialects) in the light of the notion of "phonetically empty object". The analysis proposed draws on some considerations on empty objects in Brazilian Portuguese found in Farrell (1990).

In this paper I discuss in what way the arguments of a head may be realised in a Government-Binding grammar. The particular instance of such realisation dealt with is that of the third person clitic arguments in some (colloquial) dialects of Catalan, though it is beyond the scope of this paper to give a full account of the clitic system of Catalan. In broad terms, it seems to be the case, though, that both colloquial and prescribed Catalan coincide in presenting simple sequences of clitics of first, second and third person. What is meant by simple is that each clitic can be identified in the sequence (together with epenthetic vowels which for the present purposes can be ignored). This is illustrated in (1).

- (1) a. Et veig
2 I-see
'I see you'
- b. La veig
3-F I-see
'I see her/it'

- (1) c. Me les dóna
 I 3-F-PL s/he-gives
 'S/he gives them to me'

In all these examples, each clitic can be identified with an argument of the verb, in the same way as a full NP complement; e.g. *La veig* can be a paraphrase of *Veig la sabata* 'I see the shoe' and in both sentences all the arguments of the verb *veure* 'to see', EXPERIENCER and THEME, are satisfied. This contrasts with the ungrammaticality of (2), where one of the arguments of the verb, the THEME in this case, is not realised — giving rise to a violation of the Projection Principle.

- (2) *Veig
 I-see

In common practise, the complementarity of clitics and full complements (clitic doubling aside) is expressed by postulating the following kind of structure representation for sentences with clitics:

- (3) La_i veig [pro]_i
 3-F I-see

That is, clitics satisfy the thematic requirements of their heads by coindexation with an empty object *pro* which is in argumental position (rather than being themselves in an argumental position). Nothing here depends on this mechanism; what matters is that the well-formedness of a sentence rests universally on the satisfaction of the argumental requirements of its constituents.

The hitherto regular pattern of a one-to-one correspondence between theta roles and clitics (or complements) does not apparently hold with third person clitics in some dialects of Catalan, the Barcelona one among them:

- (4) a. Li dono la carta
 [li] I-give the letter
 'I give him/her the letter'
- b. L'hi dono
 [li] I-give
 'I give it to him/her'
- (5) a. Els hi dono la carta
 [ðlzi] I-give the letter
 'I give them the letter'
- b. Els hi dono
 [ðlzi] I-give
 'I give it to them'

In these examples the clitic (or clitic sequence) *li/ðlzi* seems to stand for a DAT 3 person in (4a, 5a), and for a DAT 3 person and an ACC 3 person in (4b, 5b). Descriptions of this system can be found in the literature (e.g. Fabra (1912), Fabra (1956), Badia (1962), Wheeler (1979), Mascaró (1986)). On the understanding that it is not "extraño y bárbaro" (Fabra (1912:167)), but must conform in some way to the principles of UG, I'll suggest a syntactic account of this construction. Notice that [li] and [ðlzi] are phonetically identical in the two examples in (4) and (5) respectively (though the orthographic convention is that (4a) and (4b) are not written the same). I see no reason to expect a solution to the problem presented in the PF component of the grammar.

Let us first go through the occurrence of third person clitics systematically. (6) and (7) display the occurrences of one clitic only, (8) those of two clitics.

- (6) a. El dono a la Maria/al Joan
 3-M I-give to Maria/to Joan
 'I give it(M) to Maria/to Joan'
- b. La dono a la Maria/al Joan
 3-F
 'I give it(F) to Maria/to Joan'
- c. Els dono a la Maria/al Joan
 3-M-PL
 'I give them(M) to Maria/to Joan'
- d. Les dono a la Maria/al Joan
 3-F-PL
 'I give them(F) to Maria/to Joan'
- (7) a. [li] dono el llibre/la llibreta
 3-DAT I-give the book(M)/the booknote(F)
 'I give him/her the book/booknote'
- b. [ðlzi] dono el llibre/la llibreta
 3-PL-DAT
 'I give them the book/booknote'
- (8) a. [li] dono
 [3 DAT, 3 ACC]
- b. [ðlzi] dono
 [3 DAT, 3 PL ACC]
- b'. [lðzi] dono
 [3 DAT, 3 F PL ACC]
- c. [ðlzi] dono
 [3 PL DAT, 3 M PL ACC]

- (8) d. [ɔlzi] dono
 [3 PL DAT, 3 F PL ACC]

Some of the features that characterise the paradigms in (6) and (7) above are as follows: (a) all third person clitic clusters include [l]; (b) 3 ACC clitics may be marked as [+FEM] or [-FEM]; cf. (6a) and (6b), or (6c) and (6d). This distinction does not hold for first and second person; (c) 3 ACC clitics may be marked as [+PL] or [-PL]; cf. (6a) and (6c), or (6b) and (6d); (d) sequences including a 3 DAT are always marked by [i], and (e) 3 P DAT appear in sequences always marked by [z], i.e. as [+PL].

Does this characterisation carry through to (8), though? It is certainly true that (d) holds in (8), since all the clitic sequences stand for DATs and include an [i]. As for (a), all examples in (8) involve a third person marker [l], even though only one. With respect to (b), only one of the forms in (8) includes a [+FEM] marker: *Les hi dono (les llibretes, a ella)*. 'I give the booknotes (F) to him/her' In other cases, the distinction [+FEM]/[-FEM] is blurred. In connection with (c) and (e), i.e. for the markers for [+PL], we find one in all examples from (8b) to (8d); in (8b) and (8b') it can only correspond to the ACC argument, while in (8c) it can only correspond to the DAT. In (8d) the plural marker seems to be attributable to either DAT or ACC. The problem of attribution and distribution of markers to DAT or ACC clitics remains.

Farrell (1990) analyses sentences such as those in (9) of Brazilian Portuguese as involving an empty object.

- (9) Eu conheci *ec* numa festa
 I met *ec* at-a party

These empty objects alternate with overt full pronouns and ACC clitics, as in (10a) and (10b) respectively:

- (10) a. Eu conheci ele numa festa
b. Eu o conheci numa festa
'I met him at a party'

Farrell argues that the empty category in (9) is correctly characterised as *pro*, rather than as a variable bound by an operator *à la* Huang 1984. What concerns us in the present context is the solution to the identification problem that (9) raises. That is, in what way does the grammar guarantee that the *pro* in (9) is identified in order for (9) not to be ruled out as a violation of the Projection Principle (i.e. as a sentence missing an argument that must be saturated)? In this case, obviously, agreement on the head category is not the identificatory feature which licenses the empty object, since, unlike with empty subjects in Romance, there is no such agreement. Farrell proposes, then, to broaden the notion of identification. In his view, the intrinsic specification of features should count as identification. Accordingly, the *pro* recoverability principle of Rizzi (1986) which appears in (11) is broadened with the clause in (11').

(11) Let X be the licensing head of an occurrence of *pro*: then *pro* has the grammatical specifications of the features on X coindexed with it.

(11') Or *pro* has the grammatical specifications of designated features.

In particular, in Brazilian Portuguese, the only objects that are licensed are in the third person. First and second person pronouns must be overt in object position, as shown in (12):

- (12) Coitado do João/*de mim/*de voce. O chefe mandou *ec* embora
'Poor João/*me/*you. The boss sacked *ec*'

The licensing of *pro* in Brazilian Portuguese is possible, under (11'), by identifying it intrinsically with the feature [+3 person]. Hence, by (11'), the *ec* of Brazilian Portuguese is allowed to occur without identifying morphology. I suggest that in the problematic examples of Catalan above it is also possible to postulate an empty category with the same designated features as in Brazilian Portuguese, i.e. [+3 person]:

- (13) a. [li] *ec* dono
 [3 DAT, 3 ACC]
 b. [ðlzi] *ec* dono
 [3 PL DAT, 3 M ACC]
 c. [ðlzi] *ec* dono
 [3 PL DAT, 3 F ACC]

(The position of the empty category is irrelevant here.) Postulating such an empty category in these examples allows us to explain some phenomena otherwise unaccounted for: the lack of overt gender markers normally present in ACC forms (cf. (1)), as well as the presence of one PL marker only in (13b,c), attributed now to the DAT plural.

However, we still have to account for the distribution of this empty object of Catalan, which we will assume to be pronominal, i.e. *pro*, as in Brazilian Portuguese. Given the ungrammaticality of (2), **Veig*, we know that a structure such as that in (14) must be ruled out:

- (14) **Veig pro*
 'I-see it/her/him/them'

In actual fact, the *ec* at issue seems to occur only when a third person DAT complement has cliticised onto the verb. (I am leaving aside other possible occurrences of empty objects in Catalan

which involve generic interpretation, etc.) In general, the dialects of Catalan that we are dealing with exclude any sequence of clitics where a given clitic occurs more than once. This is not so in prescribed Catalan, where e.g. *Els els dono* 'I give them to them' is good. Suppose then that the grammar of colloquial Barcelona Catalan includes a principle excluding more than one occurrence of any given clitic in a clitic sequence. Also, this grammar includes clause (11') for the licensing of *pro*. However, unlike in Brazilian Portuguese, this *pro* only occurs when it is forced to by the "single-occurrence-of-a-clitic" filter above.

Notice, by the way, that this last filter (a filter possibly derivable from Case theory, something that I'm not going to go into here) correctly rules out sentences such as (15), where, conceivably, the two clitics *hi* and *en* could have different sources (i.e., stand for different theta roles), so that the sentences in (15) are not problematic for theta theory.

- (15) *L'hi hi dono
 *En en porta un

To summarise, the third person clitic system is constituted by the sequence of $l+\partial+z+i$, which stand respectively for [+3 person], [+FEM], [+PL], and [+DAT] (or, more accurately, a complement introduced by the preposition *a*). Cliticisation is ruled by the principles of the grammar, to which the clitic filter above is added, together with the licensing conditions on *pro*.¹

It becomes clear that the clitic system under scrutiny is consistent with a grammar which requires all argumental relationships to be recoverable.

●The author wishes to acknowledge that the research reported here is supported by a post-doctoral grant of the Spanish Ministerio de Educación y Ciencia. Any comments on it will be welcome at the author's address.

Notes

¹ The caveat must be made that a DAT singular can possibly be *l+i* (with ACC singular) or just *i* (with ACC plural), since for most speakers */*ləi*/ (DAT singular, ACC feminine singular) is out.

References

- Badia i Margarit, A. M. (1962) *Gramática catalana*, Gredos, Madrid.
- Fabra, P. (1912) *Gramática de la lengua catalana*, L'Avenç, Barcelona.
- Fabra, P. (1956) *Gramàtica catalana*, Teide, Barcelona.
- Farrell, P. (1990) 'Null Objects in Brazilian Portuguese', *Natural Language and Linguistic Theory* 8, 325-346.
- Huang, J. (1984) 'On the Distribution and Reference of Empty Pronouns', *Linguistic Inquiry* 15, 531-575.
- Mascaró, J. (1986) *Morfologia*, Enciclopèdia Catalana, Barcelona.
- Rizzi, L. (1986) 'Null Objects in Italian and the Theory of *pro*', *Linguistic Inquiry* 9, 507-557.
- Wheeler, M. (1979) *Phonology of Catalan*, Basil Blackwell, Oxford.

Algemene Taalwetenschap

Universiteit van Amsterdam

Spuistraat 210

NL-1012 VT Amsterdam