

Case-theory: a solution of the bound pronoun problem in Romance languages

JELLY JULIA DE JONG

1. Bound pronouns o.k.

The standard binding conditions A and B as formulated in Chomsky (1981)

- (1) A) An anaphor has to be bound in its governing category
 - B) A pronoun has to be free in its governing category
- (cf. *Lectures on Government and Binding*, p.188)

are not able to account for the so-called snake-sentences:

- (2) John_i saw a snake near him_i

because in this case the pronoun *him* is bound in its governing category (GC).

This is in conflict with principle B of the standard binding conditions which states that a pronoun has to be free in its GC. Empirical investigation has shown that in several Romance languages we find bound pronouns in snake-sentences also. We can see this in (3), (4), (5):

- (3) Spanish: Juan vio una serpiente cerca de él.
'John saw a snake near him'
- (5) Portuguese: O Joao viu um gato ao pé dele.
'John saw a cat near him'
- (4) Catalan: En Joan va veure una serp a prop d'ell.
'John saw a snake near him'

Lees and Klima (1963) proposed a sentential solution for (2):

- (6) John saw [a snake to be near him]

This is what we would call a small clause solution now. Chomsky (1981) briefly considers such solutions in chapter 5 but rejects them as being not generally applicable to all cases. As shown by Koster (1985), the inadequacy of the small clause solution to the problem clearly appears in Dutch cases with intransitives, as we can observe in (7):

- (7) Jan keek om zich heen.
'John looked around him'

If the domain introduced by *om* were a small clause, we would have the structure as in (8):

- (8) Jani keek [_S PRO_i om zich_i heen]

The problem is that with such intransitives, there is no other antecedent for PRO than the subject itself. But since *zich* is also bound by the subject, the reflexive is bound in the domain in question, while it should be free. In other words, with a structure like (8) one would expect *zichzelf* rather than *zich*. Thus, the small clause analysis gives exactly the opposite results from what is needed. This analysis must therefore be rejected.

Moreover, several Romance languages allow pronouns to be bound in more local contexts, as we can observe in (9), (10) and (11):

- | | |
|--|--|
| <p>(9) Spanish:</p> <p>a. Juan habla de él.</p> <p>b. Juan habla de él mismo.</p> <p>c. Juan habla de sí mismo.
'John talks about himself'</p> | <p>(10) Catalan:</p> <p>a. En Joan parla dell.</p> <p>b. En Joan parla dell mateix.</p> <p>c. En Joan parla de si mateix.
'John talks about himself'</p> |
| <p>(11) Portuguese:</p> <p>a. O Joao fala sempre dele.</p> <p>b. O Joao fala sempre dele próprio / dele mesmo.</p> <p>c. O Joao fala sempre de si próprio / de si mesmo.
'John talks always about himself'</p> | |

As we can observe, besides the (b) and (c) sentences which contain a SELF-anaphor and a combination of a SE- and a SELF-anaphor respectively,¹ also the (a) variants are grammatical. Assuming that *él*, *ell* and *ele* are pronouns, the standard binding conditions are not able to account for these locally bound pronouns.

2. Bound pronouns not o.k.

We have seen briefly in which kind of context a bound pronoun is allowed. Now let us study a context in which a bound pronoun is not acceptable.

- | | |
|---|---|
| <p>(12) Spanish:</p> <p>a. *Juan se ama a él.</p> <p>b. *Juan se ama a él mismo.</p> <p>c. Juan se ama a sí mismo.
'John loves himself'</p> | <p>(13) Catalan:</p> <p>a. *En Joan sestima a ell.</p> <p>b. ?En Joan sestima a ell mateix.</p> <p>c. En Joan sestima a si mateix.
'John loves himself'</p> |
| <p>(14) Portuguese:</p> <p>a. *O Joao ama-se a ele.</p> <p>b. O Joao ama-se a ele próprio / ?a ele mesmo.</p> <p>c. O Joao ama-se a si próprio / a si mesmo.
'John loves himself'</p> | |

(1) Reinhart and Reuland (1989) introduce the terms SE- and SELF-anaphors. Instantiations of SE-anaphors are e.g. *zich* (Dutch), *seg* (Norwegian). Instantiations of SELF-anaphors are e.g. *zelf* (Dutch), *self* (English).

The (a) examples are ungrammatical because a bound pronoun is not allowed here. As we can observe in the (c) sentences, the examples can be saved by substituting the pronoun by a combination of a SE- + SELF-anaphor. Further investigation should clarify why the (b) variants, with a combination of a pronoun and a SELF-anaphor, are ungrammatical in Spanish, dubious in Portuguese and correct for some Catalan speakers and dubious or highly dubious for other Catalans. This question would be much easier to explain if the (b) variants would be definitely wrong in all three languages mentioned. We then could claim that *mismo* is just an emphatic element which has no status of anaphor whatsoever. The (b) sentences, then, would all be instances of bound pronouns, just like the (a) sentences.

3. Chain conditions

The problem we want to tackle here is the problem of the pronouns that in Ibero-Romance languages are locally bound. The anaphoric system of Frisian presents a serious problem for the standard binding theory too. Like English, Frisian has a two-member system; there is an anaphor *himsels*, and a pronominal *him*. However, unlike Dutch and English, Frisian has locally bound pronominals. The generalization is that wherever Dutch allows *zich* (a so-called SE-anaphor), Frisian allows a bound pronominal. Two pronominals, namely the 3rd person singular feminine and the 3rd person plural (common gender) have two object forms: both have *se* as well as *har* (or —in plural— *harren*). Often, they are used interchangeably. However, unlike *har / harren*, *se* is ungrammatical when locally bound. J. Hoekstra (1991) shows that *se* is ungrammatical in the next positions: in the object position of prepositions (which assign oblique Case), in the object position of transitive adjectives (which assign oblique Case to their objects, cf. Van Riemsdijk, 1983), in the experiencer argument position of psychological verbs (As Den Besten 1984 and also Belletti & Rizzi 1988 show, this position is assigned oblique Case), in free dative constructions (which are assigned inherent dative Case). Hoekstra comes to the conclusion that '*Se* must bear structural Case'. Since he shows that in all four circumstances just mentioned the pronoun *har(ren)* is allowed, there is another, implicit, conclusion: '*Har(ren)* can bear inherent Case'. There is a fifth circumstance where *se* is ungrammatical and *har(ren)* is grammatical, namely in a bound position in a reflexive construction. To account for these binding facts in Frisian Reinhart and Reuland (1991a) have developed a Generalized Chain condition on A-chains which is based on a general notion of Chain links by Chomsky (1986a, b):

(15) Generalized Chain condition

$C = (\alpha_1, \dots, \alpha_n)$ is a chain iff C is the maximal sequence such that conditions a and b are fulfilled:

- a) there is an index i such that for all j , $1 \leq j \leq n$, α_j carries that index
- b) for all j , $1 \leq j \leq n$, α_j governs α_{j+1}
(Reinhart and Reuland 1991a)

In a sentence like

- (16) a. *Jeltsje skammet har.* b. **Jeltsje skammet se.* 'Julia is ashamed of herself

the pronoun *se* which bears structural Case produces an ungrammatical sentence when it appears in a bound position (16b) but *har* is o.k. Reinhart and Reuland argue that the reason why (16a) is grammatical is because *har* bears inherent Case and that in the entire sentence there is only one element which is specified for structural Case: the proper name *Jeltsje*. They then propose a revision of the Chain condition:

(17) Revision of the Generalized Condition on A-chains:

A maximal A-chain ($\alpha_1, \dots, \alpha_n$) contains precisely one link - α_1 - which is fully specified for grammatical features (structural Case features and phi-features)

They state that pronominals are fully specified for phi-features. The referential dependence of anaphors is syntactically reflected in having a paradigm which lacks a distinction in at least one grammatical dimension. Anaphors fail to have a full paradigm for singular-plural or gender distinctions (Reinhart & Reuland 1991b).

Thus, according to this revised condition on chains, *Jeltsje* in (16a) forms a chain with *har*. On the other hand, (16b) contains two elements bearing structural Case and thus violates the Chain condition in (17). We will see that this Chain condition accounts for the (un)grammaticality of locally bound pronouns in Spanish, Portuguese and Catalan.

- | | |
|---|--|
| (18) Pedro habla de él.
'Peter talks about himself' | (20) Gregorio lo ha comprado para él.
'Gregory it has bought for himself' |
| (19) Javier ha construido la casa de la playa para él.
'Xavier has built the house on the beach for himself' | |
| (21) a. *Mario se ama a él.
b. Mario se ama a sí mismo.
'Mario loves himself' | (22) a. *Fernando se lava a él.
b. Fernando se lava a sí mismo.
'Fernando washes him/self' |

The prepositions in (18)-(20) assign oblique Case to the object. This means that these objects do not have structural Case. The only element in these sentences which is fully specified for all grammatical features is the subject. This means the objects are able to form a Chain with the subject without violating the Chain Condition (17). But what about (21) and (22)? Why can't we have bound pronouns in the (a) sentences?

4. Preposition *a* as a dummy Case-marker

On first sight all object pronouns in (18) - (22) are assigned oblique, thus inherent, Case by the preceding preposition. But the value of the preposition is not identical in all sentences. In Spanish, Portuguese and Catalan indirect objects and strong direct object pronouns with the features [+Animate] and [+Specific] are preceded by the preposition *a*. For example:

- | | |
|---|---|
| (23) Spanish:
Me ha visto a mí.
'(me) he-has-seen me' | (24) Catalan:
Jo el corregiré a ell.
'I (him) will-correct him' |
|---|---|

- (25) Portuguese:
 O homem burlou-te a ti.
 'The man has-cheated (you) you'

Jaeggli (1982) points out that the preposition *a*, when it is inserted before a strong direct object pronoun with the features ([+Animate], [+Specific]), is a dummy Case-marker and he was the first to connect a Case assignment mechanism to the phenomenon of clitic-doubling. Let us consider the circumstances in which clitic-doubling is allowed in Spanish:

- | | |
|----------------------------------|----------------------------|
| (26) a. Veo a los chicos. | (27) a. Compró los libros. |
| b. Los veo a ellos. (los chicos) | b. *Los compró los libros. |
| 'I see them' (the boys)' | 'I buy the books' |

In (27b) the object *libros* (books) is not [+Animate] so no *a*-insertion takes place. In (27b) the clitic absorbs Case, this means Case cannot be assigned to the object *libros*. In (26b) the clitic also absorbs Case but insertion of the preposition *a* takes place because the object is [+Animate] and [+Specific]. As this preposition *a* is a dummy Case-marker it is able to assign Case to the object *ellos*. This is the reason why (26b) is grammatical and (27b) is ungrammatical. The dummy Case-marking character of the preposition *a* is reason for Demonte (1987) to suggest to make a distinction between true PPs and pseudo-PPs (=NPs) in the grammar of Romance languages. In her reasoning the *a* + NP sequences of (21)-(26) are pseudo-PPs.

Now, let us have a look again at sentences (21) and (22) and let us see whether we have a solution for the ungrammaticality of the (a) sentences.

- | | |
|-----------------------------|---------------------------------|
| (21) a. *Mario se ama a él. | (22) a. *Fernando se lava a él. |
| b. Mario se ama a sí mismo. | b. Fernando se lava a sí mismo. |
| 'Mario loves himself' | 'Fernando washes him/self' |

We now assume the standard idea that the preposition *a* in the sentences (21)-(25) is a dummy Case-marking preposition whereas the head of the PP in the sentences (18)-(20) is a 'real' preposition with 'true' semantic value. In the sentences (21) and (22) the preposition *a* is a dummy Case-marking preposition which assigns structural Case (accusative) to the object. This means that in these sentences we find two elements bearing structural Case, violating thus the Chain Condition (17). We can observe in (21) and (22) that indeed the pronoun produces an ungrammatical sentence, whereas the anaphor *sí mismo* is allowed.

The next question is: why is there obligatory insertion of the preposition *a* in (26c)?

- (26) c. *Veo los chicos.

Why do we need the dummy Case-marker *a* in (26a) when the verb can assign Case to the object NP? Let us consider the next sentences:

- | | |
|--------------------|--------------------|
| (28) Veo al chico. | (29) Veo un chico. |
| 'I see the boy' | 'I see a boy' |

- (30) Veo a un chico que es profesor.
 'I see a boy who is a teacher'

We can observe that the object NPs in (28) and (30) are [+Specific]: in (28) it is the definite article that specifies the NP; in (30) it is a relative clause which modifies the NP. In both cases *a*-insertion takes place, even when the object is preceded by an indefinite article. We must come to the conclusion that [+Specific] [+Animate] direct objects need a different objective Case which is assigned to it by the dummy Case-marking preposition. The next sentences give some more evidence for this hypothesis:

- (31) a. Busco a una secretaria que está vestida de blanco.
 I-am-looking-for a secretary who is (IND) dressed in white
 b. Busco una secretaria que sepa hablar inglés.
 I-am-looking-for a secretary who speaks (SUBJ) English

As we can observe, in sentence (31a) *a*-insertion takes place because the NP *secretaria* is modified by a relative clause. However, in sentence (31b) no *a*-insertion takes place notwithstanding the fact that the NP is modified by a relative clause. The use of the subjunctive mood of the verb in the relative clause implies that the object it modifies is a non-specific NP (cf. (31b), despite of its modification by the relative clause. It is the kind of sentence one would expect to find in written form in an add in a newspaper: Company is looking for any person that is able to function as a secretary and that is able to speak English. So we can say the object NP in (31b) is non-specific. The use of the indicative mood in (31a) already indicates that the object NP in this sentence is specific. One expects to hear this sentence in spoken form expressed by someone who is inquiring after the whereabouts of a certain person describing her as a secretary dressed in white.

5. Two types of objective Case

In the work of various authors we can find the idea, put forward here for Spanish, that there is a correlation between the type of interpretation an object gets and the type of Case assigned to this object. Belletti (1988) notes that in Finnish there are two possible Cases for an object NP. Depending on the reading associated with it, the object of a transitive verb will be marked either with accusative or with partitive Case. De Hoop (1992) notes that Finnish is not the only language that shows morphological realization of two different objective Cases. In Turkish direct objects optionally get an accusative Case-marker. Once again we can observe here a correlation between the type of objective Case and the interpretation of the NP. In Turkish inherent Case correlates with non-specificity whereas structural Case corresponds to specificity (cf. Enç 1991). In Greenlandic Eskimo there also is a correlation between the type of Case assigned to an object and the type of interpretation this object receives. An indefinite object NP in this language has instrumental Case in an antipassive construction and a definite object NP has nominative Case in a transitive variant (cf. Bittner 1988 and Bok-Bennema 1991). De Hoop puts forward that an object gets a *strong* reading if and only if it bears *strong* structural Case. A 'strong'

reading of an object is attributed to one certain type of NP, viz. the type of a generalized quantifier; the term 'strong reading' is meant to capture the unmarked reading of strong NPs as well as strong readings of weak NPs such as referential (specific), partitive, and generic readings. It appears to be important to talk about 'strong' and 'weak' readings of objects instead of '(in)definite' objects since, as our Spanish data also show, we have to distinguish (at least) two classes of indefinite objects. Furthermore, within De Hoops approach, the two types of objective Case that are distinguished are both claimed to be structural rather than inherent.² The type of structural Case that is related to the strong reading of an object is called *strong* Case and is licensed at S-structure, whereas the other type of structural objective Case is assigned at D-structure and is called *weak* Case. This seems to be in accordance with the Spanish facts. It is in accordance, too, with Zubizarreta (1985).

Zubizarreta (1985) shows, like Jaeggli (1982), that *a* is a dummy preposition functioning as a semantically empty Case-marker by pointing out that the accusative direct object preceded by *a* is semantically unrestricted. She argues that an argument that is realized as object of the verb or as subject is semantically unrestricted since other roles than Agent may be assigned to the subject position and other roles than Theme may be assigned to the object position. However, the semantic role of an argument that is realized in a prepositional phrase is restricted by the preposition: the object of *to* must be the Goal, the object of *from* must be the Source, the object of *in* must be a Location, etc. She then shows in the examples mentioned here under (32) that the *a* preceding the animate direct object does not semantically restrict the object. Thus, she concludes, *a* is a semantically empty Case-marker.

- (32) a. Juan *la* trajo *a* María a casa. b. El mucamo *la* sirvió *a* María.
 'John brought María home'. 'The maid served María'
 c. El mucamo *le* sirvió la comida *a* María.
 'The maid served María the food'.

In (32a) the accusative direct object *María* is a Theme. In (32b) it is a Goal, comparable to the dative indirect object in (32c). This implies that *a* assigns structural Case to its object rather than inherent Case. It seems very likely that the other type of objective Case which gives rise to a weak reading, is also a structural Case.

6. Conclusion

In the previous section I have been using the terms 'specific' and 'non-specific' NP's. I would like to explain very briefly what I mean by these terms. I will not give an extensive exposition of the theory that describes these terms and its explanations nor will I go into any details concerning the discussion on these topics that is taking place at the moment. I just would like to outline in a rather intuitive way what I mean by a specific NP and by a 'non-specific' NP. Let us consider the next sentence:

- (33) A colleague of mine went nuts (because of linguistics).

(2) Structural Case is assigned in certain configurations, whereas inherent Case is related to a specific θ -role (cf. Chomsky 1986; Belletti 1988).

Besides the existential reading, the indefinite NP in (33) also has the interpretation of a referring expression, comparable to the interpretation of a proper name or a demonstrative. That is, the existential reading just states that the set of crazy colleagues of mine is not empty, whereas the referential reading refers to a specific individual, for instance {X}. The existential reading is also called non-specific in the literature, whereas the referential one is called 'specific'.

Fodor and Sag (1982), for example, present evidence for a semantic ambiguity in indefinite NPs. They provide a number of arguments in favour of the hypothesis that the difference between an existential and a referential reading for indefinites reflects a semantic ambiguity, over and above contextual scope ambiguities. They sum up certain factors that favour either a referential or an existential reading of an indefinite NP. When we observe our Spanish data we come to the conclusion that the next factor can be included among those mentioned by Fodor and Sag:

(34) n. For Spanish:

The use of the preposition *a* before the object and the use of the indicative mood in the relative clause that modifies the object gives rise to a referential reading whereas the omission of this preposition and the use of the subjunctive mood leads to an existential reading of the indefinite NP.

According to Jaeggli (1982) the difference between specific NP's, which are *a*-NP's, and non-specific NP's which are 'bare' NP's is related to Case. We saw this is indeed plausible since according to Bittner (1988), Bok-Bennema (1991), Enç (1991) and De Hoop (1992) the type of interpretation an object gets is correlated to the type of Case assigned to this object. Following De Hoop (1992), on a link between two types of structural Case and different readings on objects, it can be argued that Spanish *a* is a Case-marker inserted at S-structure in order to license strong structural Case on [+Specific] [+Animate] objects. In the case of Spanish, it seems very likely that the other type of objective Case which gives rise to a weak reading, is also a structural rather than an inherent Case. Furthermore, Jaeggli's hypothesis is also in accordance with the Chain Condition developed by Reinhart and Reuland (1991a,b). In fact, Reinhart and Reuland's A-chain condition can be adapted somewhat in the sense that the notion *structural Case* is replaced by *strong structural Case* in order to distinguish the latter type of Case from inherent as well as from weak structural Case with respect to A-chains (this was independently put forward by De Hoop as well).

References

- Belletti, A., 1988, "The Case of unaccusatives", *LI* 19: 1, 1-34.
 ——— and Rizzi, L., 1988, "Psych-verbs and θ -theory", *NLLT* 6, 291-352.
 Besten, H. den, 1984, "The ergative hypothesis and free word order in Dutch and German" in Toman J. (ed.); *Studies in German grammar*; Studies in Generative Grammar, n. 21; Foris, Dordrecht, 23-64.
 Bittner, M., 1988, *Canonical and noncanonical argument expressions*; doctoral dissertation, University of Texas, Austin.

- Bok-Bennema, R., 1991, *Case and agreement in Inuit*. Studies in Generative Grammar, n. 38; Foris, Dordrecht.
- Chomsky, N., 1981, *Lectures on government and binding. The Pisa lectures*, Foris Publications, Dordrecht.
- , 1986a, *Knowledge of Language: Its nature, origin and use*; Convergence, Praeger.
- , 1986b, *Barriers*. Linguistic Inquiry Monographs, n 13, MIT Press, Cambridge.
- Demonte, V., 1987, "C-command, prepositions, and predication", *LI* 18.1, 147-157.
- Enç, M., 1991, 'The semantics of specificity', *LI* 22.1, 1-26.
- Fodor, J. and Sag, I., 1982, "Referential and quantificational indefinites", *Linguistics and Philosophy* 5., 355-398.
- Hoekstra, J., 1991, *Pronouns and Case. On the distribution of Frisian harren and se them and the relation between abstract Case and morphological case*; ms. Fryske Akademy, Ljouwert.
- Hoop, H. de, 1992, *Case configuration and noun phrase interpretation*. Doctoral dissertation, University of Groningen.
- Jaeggli, O., 1982, *Topics in Romance Syntax*. Studies in Generative Grammar, n. 12; Foris, Dordrecht.
- Koster, J., 1985, *Reflexives in Dutch. Grammatical Representation*. Studies in Generative Grammar, n. 22, 141-167. Foris, Dordrecht.
- Lees, R. and Klima, E., 1963, "Rules for English pronominalization", *Lg* 39.1, 17-28.
- Reinhart, T. and Reuland, E., 1989, "Anaphors and logophors: An argument structure perspective", in J. Koster, E. Reuland (eds.) *Long-distance anaphora*, Cambridge University Press, Cambridge, 1991, p. 283-321.
- , 1991a, "Pronouns, simplex anaphors and chains". *The Seventh Workshop on Comparative Germanic Syntax*, November 22 - 24, 1991.
- , 1991, *Binding conditions and chains*; ms. University of Tel Aviv / Groningen.
- Riemsdijk, H. van, 1983, "The Case of German adjectives" in: Heny, F. and Richards, B., , *Linguistic categories: auxiliaries and related puzzles. Vol. 1: Categories*; Dordrecht. Reidel, 223-252.
- Zubizarreta, M. L., 1985, "Morphophonology and morphosyntax: Romance causatives", *LI* 16.2, 247-289.